

STIC-Biotech/ChemLib

From: Stucker, Jeffrey
Sent: Tuesday, September 09, 2003 8:39 AM
To: STIC-Biotech/ChemLib
Subject: 09/980777

I need to search the art for resistance conveying mutations in HIV-2 protease, specifically at position 90. Unfortunately, my application, 09/980777, does not give the complete sequence of HIV-2 protease. Can you use a publicly available sequence to run this search? The most desirable (but not only) mutation is L90M.

Thanks,
Jeffrey Stucker
1648
CM1 8E01
703-308-4237
Mailbox: 8E12

Searcher: D. Schreiber
Phone: 308-4237
Location: CM1 6A03
Date Picked Up: _____
Date Completed: 9/25
Searcher Prep/Review: 9/25
Clerical: 9/25
Online time: 19 91

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: 4
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: CompuGen
WWW/Internet: _____
Other (specify): _____

24

THIS PAGE BLANK (USPTO)



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 104576

**TO: Jeffrey Stucker
Location: CM1/8E01/8E12
Art Unit: 1648
Friday, September 26, 2003**

Case Serial Number: 09/980777

**From: David Schreiber
Location: Biotech-Chem Library
CM1-6A03
Phone: 308-4292**

david.schreiber@uspto.gov

Search Notes

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 25, 2003, 11:52:31 ; Search time 50 Seconds
(without alignments)
3288.815 Million cell updates/sec

Title: POL_HV2RO

Perfect score:

Sequence: 1 TGRFFRTGPGKEAPQLPRG.....RQEMDSGSHLEGAREDEGEMA 1036

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Minimum DB seq	length: 0
Maximum DB seq	length: 2000000000

Post-processing: Minimum Match 0%

Fast processing. Minimum Match 0%
Maximum Match 99%

Maximum match 55%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*

1:	/SIDSL/cgcdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2:	/SIDSL/cgcdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3:	/SIDSL/cgcdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4:	/SIDSL/cgcdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5:	/SIDSL/cgcdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6:	/SIDSL/cgcdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7:	/SIDSL/cgcdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8:	/SIDSL/cgcdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9:	/SIDSL/cgcdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10:	/SIDSL/cgcdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11:	/SIDSL/cgcdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12:	/SIDSL/cgcdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13:	/SIDSL/cgcdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14:	/SIDSL/cgcdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15:	/SIDSL/cgcdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16:	/SIDSL/cgcdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17:	/SIDSL/cgcdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18:	/SIDSL/cgcdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19:	/SIDSL/cgcdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20:	/SIDSL/cgcdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21:	/SIDSL/cgcdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22:	/SIDSL/cgcdata/geneseq/geneseqp-emb1/AA2001.DAT.*
23:	/SIDSL/cgcdata/geneseq/geneseqp-emb1/AA2002.DAT.*
24:	/SIDSL/cgcdata/geneseq/geneseqp-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	5426	99.5		1036	13	AAR20599	ROD HIV-2 polymera
2	5426	99.5		1036	21	AAV51978	HIV-2 ROD isolate
3	5378.5	98.7		1027	9	AAP81773	Sequence encoded b
4	5310	97.4		1014	9	AAP80810	Sequence of pol pr
5	5075.5	93.1		3080	10	AAP93285	Sequence of clone
6	5059	92.8		3210	9	AAP81771	Deduced sequence e
7	5055.5	92.7		1055	18	AAW13055	HIV-2 provirus-enc
8	5023.5	92.1		1035	11	AAR404025	Pol gene product o
9	4629	84.9		1060	20	AAW899314	SIVmac239 genome p

10	4624	84.8	1056	9	AAP81783	Sequence encoded b
11	4616	84.7	1060	13	AAAP23366	siVmac239 pol gene
12	4598	84.3	1056	9	AAAP80809	Sequence of pol pr
13	4519.5	82.9	1055	21	AAB12993	HIV-2 reverse tran
14	3324	61.0	1018	21	AAB01769	Simian immunodef
15	3296	60.5	1016	22	AAAG65228	Mandrill immunode
16	3157	57.9	1002	9	AAAP81861	Sequence encoded b
17	3157	57.9	1002	19	AAAP72993	HIV isolate LAV, MA
18	3150.5	57.8	1014	19	AAW68474	HIV-1 strain YBF30
19	3134.5	57.5	1005	21	AAAB69287	HIV-1 non-subtype
20	3133	57.5	1003	21	AAAB69288	HIV-1 non-subtype
21	3132	57.4	1003	11	AAAR0060	HIV-1 pol protein
22	3127.5	57.4	1000	21	AAAB69282	HIV-1 non-subtype
23	3120	57.2	1002	21	AAAB69283	HIV-1 non-subtype
24	3119	57.2	1003	7	AAAP61508	Sequence of ARV-2
25	3119	57.2	1003	21	AAV77301	HIV-1 (ATCC CRL 85
26	3117	57.2	1003	21	AAV70601	Corrected version
27	3117	57.2	1003	21	AAV70602	Codon optimised Hu
28	3116.5	57.2	1003	21	AAAB69284	HIV-1 non-subtype
29	3116	57.2	1003	13	AAAR29705	pol gene decoded f
30	3115	57.1	1003	8	AAAP70861	Sequence encoded b
31	3114	57.1	1015	14	AAAR43875	HTLV-III POL gene
32	3113	57.1	1012	7	AAAP61507	Sequence of revers
33	3113	57.1	1012	20	AAW89323	HIV-1 pol protein
34	3112	57.1	1003	21	AAAB10047	HIV-1 pol protein,
35	3112	57.1	1003	21	AAV70600	Wild type Human Im
36	3112	57.1	1015	7	AAAP60347	HTLV-III virus (HI
37	3112	57.1	1015	14	AAAR43867	HTLV-III POL gene
38	3112	57.1	1015	22	AAAB85993	Amino acid sequenc
39	3112	57.1	3025	22	AAAB86169	HIV-1 subtype C pr
40	3111	57.1	1003	23	AAO19387	Lymphadenopathy-as
41	3110	57.0	1003	7	AAAP60420	Sequence of LAV v1
42	3110	57.0	1003	21	AAAB69279	HIV-1 non-subtype
43	3109	57.0	1003	24	AAUB5752	AIDS associated re
44	3108.5	57.0	995	23	AAU84391	HIV POL consensus
45	3104.5	56.9	1012	20	AAW90176	HTLV-III pol prote

ALIGNMENTS

RESULT 1

AAR20599

ID AAR20599

XX

AC AAR20599;

XX

DT	25-MAR-2003	(updated)
25	25-MAR-2003	25-MAR-2003
26	26-MAR-2003	26-MAR-2003
27	27-MAR-2003	27-MAR-2003
28	28-MAR-2003	28-MAR-2003
29	29-MAR-2003	29-MAR-2003
30	30-MAR-2003	30-MAR-2003
31	31-MAR-2003	31-MAR-2003
32	31-MAR-2003	31-MAR-2003
33	31-MAR-2003	31-MAR-2003
34	31-MAR-2003	31-MAR-2003
35	31-MAR-2003	31-MAR-2003
36	31-MAR-2003	31-MAR-2003
37	31-MAR-2003	31-MAR-2003
38	31-MAR-2003	31-MAR-2003
39	31-MAR-2003	31-MAR-2003
40	31-MAR-2003	31-MAR-2003
41	31-MAR-2003	31-MAR-2003
42	31-MAR-2003	31-MAR-2003
43	31-MAR-2003	31-MAR-2003
44	31-MAR-2003	31-MAR-2003
45	31-MAR-2003	31-MAR-2003
46	31-MAR-2003	31-MAR-2003
47	31-MAR-2003	31-MAR-2003
48	31-MAR-2003	31-MAR-2003
49	31-MAR-2003	31-MAR-2003
50	31-MAR-2003	31-MAR-2003
51	31-MAR-2003	31-MAR-2003
52	31-MAR-2003	31-MAR-2003
53	31-MAR-2003	31-MAR-2003
54	31-MAR-2003	31-MAR-2003
55	31-MAR-2003	31-MAR-2003
56	31-MAR-2003	31-MAR-2003
57	31-MAR-2003	31-MAR-2003
58	31-MAR-2003	31-MAR-2003
59	31-MAR-2003	31-MAR-2003
60	31-MAR-2003	31-MAR-2003
61	31-MAR-2003	31-MAR-2003
62	31-MAR-2003	31-MAR-2003
63	31-MAR-2003	31-MAR-2003
64	31-MAR-2003	31-MAR-2003
65	31-MAR-2003	31-MAR-2003
66	31-MAR-2003	31-MAR-2003
67	31-MAR-2003	31-MAR-2003
68	31-MAR-2003	31-MAR-2003
69	31-MAR-2003	31-MAR-2003
70	31-MAR-2003	31-MAR-2003
71	31-MAR-2003	31-MAR-2003
72	31-MAR-2003	31-MAR-2003
73	31-MAR-2003	31-MAR-2003
74	31-MAR-2003	31-MAR-2003
75	31-MAR-2003	31-MAR-2003
76	31-MAR-2003	31-MAR-2003
77	31-MAR-2003	31-MAR-2003
78	31-MAR-2003	31-MAR-2003
79	31-MAR-2003	31-MAR-2003
80	31-MAR-2003	31-MAR-2003
81	31-MAR-2003	31-MAR-2003
82	31-MAR-2003	31-MAR-2003
83	31-MAR-2003	31-MAR-2003
84	31-MAR-2003	31-MAR-2003
85	31-MAR-2003	31-MAR-2003
86	31-MAR-2003	31-MAR-2003
87	31-MAR-2003	31-MAR-2003
88	31-MAR-2003	31-MAR-2003
89	31-MAR-2003	31-MAR-2003
90	31-MAR-2003	31-MAR-2003
91	31-MAR-2003	31-MAR-2003
92	31-MAR-2003	31-MAR-2003
93	31-MAR-2003	31-MAR-2003
94	31-MAR-2003	31-MAR-2003
95	31-MAR-2003	31-MAR-2003
96	31-MAR-2003	31-MAR-2003
97	31-MAR-2003	31-MAR-2003
98	31-MAR-2003	31-MAR-2003
99	31-MAR-2003	31-MAR-2003
100	31-MAR-2003	31-MAR-2003

DT 19-MAY-1992 (first entry)

XX
DE
POD UTY-3 52DE
XX
KOD HIV-2 po

Human immunodeficiency virus

KW
XX
HUMAN IMMUNODEFICIENCY VIRUS

Human immunodeficiency

XX
03
HUMAN IMMUNODEFICIENCY VIRUS-2 ROD ISOLATE.

PN US5079342-A.

XX
XX

PD 07-JAN-1992.

XX
XX

XX	PS	Disclosure; Page 13; 30pp; English.
XX	CC	The amino acid sequence is that of a polymerase from the ROD HIV-2 isolate, it is encoded by nucleotides 1829-4936. See also AAR20596-CC AAR20596-R20608 and AAR2745-R2749.
XX	CC	(updated on 25-MAR-2003 to correct PF field.)
XX	SQ	Sequence 1036 AA;
Query Match 99.5%; Score 5426; DB 13; Length 1036;		
Best Local Similarity 99.5%; Pred. No. 0;		
Matches 1031; Conservative 1; Mismatches 4; Indels 0; Gaps 0;		
QY	1	TGRFFRTGPGKEAPOLPGPSAGADTNTSPGSSSGSTGEIYAAREKTERAERETIQ 60
DB	1	TGRFFRTGPGKEAPOLPGPSAGADTNTSPGSSSGSTGEIYAAREKTERAERETIQ 60
QY	61	SDRGLTAPRAGGDTIQGATNRGLAAPQFSLWKRPPVTVAYIEGQPVVELLDTGADDSIVAG 120
DB	61	SDRGLTAPRAGGDTIQGATNRGLAAPQFSLWKRPPVTVAYIEGQPVVELLDTGADDSIVAG 120
QY	121	IELGNNYSKPIVGGIGGFTINTEYKNVEIEVLNKKVRATIMTGDTPINIFGRNILTALGM 180
DB	121	IELGNNYSKPIVGGIGGFTINTEYKNVEIEVLNKKVRATIMTGDTPINIFGRNILTALGM 180
QY	181	SLNLPVAKVEPIKIMLPKGDGPKLRQWPLTKREKTEALKEICEKMEKEGQLEEAAPTNPY 240
DB	181	SLNLPVAKVEPIKIMLPKGDGPKLRQWPLTKREKTEALKEICEKMEKEGQLEEAAPTNPY 240
QY	241	NPTFAIKKKDKNKRMLIDFRELNVKQTDFTIEIQLGIPHPAGLAKKRITVLDVGDAYF 300
DB	241	NPTFAIKKKDKNKRMLIDFRELNVKQTDFTIEIQLGIPHPAGLAKKRITVLDVGDAYF 300
QY	301	STPLHEDFPYTAFTLPSVNNAPGKRYIKVLPQGWKGSIPAIFQHTMRQVLEPRKANK 360
DB	301	STPLHEDFPYTAFTLPSVNNAPGKRYIKVLPQGWKGSIPAIFQHTMRQVLEPRKANK 360
QY	361	DVIIIOYDDIIASDRTDLEHDRVVLQKLLNGLGFSTPDEKFKQDPYPYHWMGYELWP 420
DB	361	DVIIIOYDDIIASDRTDLEHDRVVLQKLLNGLGFSTPDEKFKQDPYPYHWMGYELWP 420
QY	421	TKWKLOKIQLPQKEIWTVNDIQKLVGLNWAQLYPGIKTKHLCLRLIRKMTLITEVQWT 480
DB	421	TKWKLOKIQLPQKEIWTVNDIQKLVGLNWAQLYPGIKTKHLCLRLIRKMTLITEVQWT 480
QY	481	ELAEAELENRIILSQEQGHYQYQEKLEATVQKDOENQWYTKIHQBEKILKVGKYAKV 540
DB	481	ELAEAELENRIILSQEQGHYQYQEKLEATVQKDOENQWYTKIHQBEKILKVGKYAKV 540
QY	541	KNHTNGIRLLAQVQKIGKEALVTWGRIPKPHLPVEREINQWQNDYQWTVIPDWFV 600
DB	541	KNHTNGIRLLAQVQKIGKEALVTWGRIPKPHLPVEREINQWQNDYQWTVIPDWFV 600
QY	601	STPPLVRLAFNLVGDPIPGAETFTYDGGSCNRQSKGKAGYVTDGKDKVKKLEQTTNQA 660
DB	601	STPPLVRLAFNLVGDPIPGAETFTYDGGSCNRQSKGKAGYVTDGKDKVKKLEQTTNQA 660
QY	661	ELEAFAMALTDGSPKVNIIQVDSQYVNGISASOPTSESKIVNQIIEEMIKKEAIVVAVNP 720
DB	661	ELEAFAMALTDGSPKVNIIQVDSQYVNGISASOPTSESKIVNQIIEEMIKKEAIVVAVNP 720
QY	721	AHKGIGGQEVHLYSQGIRQVLFLEKTEPAQEEHEKYHSNVKELSHKFGIPNLVARQIV 780
DB	721	AHKGIGGQEVHLYSQGIRQVLFLEKTEPAQEEHEKYHSNVKELSHKFGIPNLVARQIV 780
QY	781	NSCAOQQKGEAIGHGVNAELGTWQMDCTHLECKIIIVAVHVASGFIEAIVPQSGRQT 840
DB	781	NSCAOQQKGEAIGHGVNAELGTWQMDCTHLECKIIIVAVHVASGFIEAIVPQSGRQT 840
QY	841	ALFLLKLSRWPIITHLTDNGANFTSQEVKVAWVWIGIEQSGVPYNPQSGVQVWAMNHH 900
DB	841	ALFLLKLSRWPIITHLTDNGANFTSQEVKVAWVWIGIEQSGVPYNPQSGVQVWAMNHH 900

QY	901	LKNQISRIREQANTTETIVLMAIHCNFKRRGGIGDMTPSERLINMITTEQEIQLQAKN 960
DB	901	LKNQISRIREQANTTETIVLMAIHCNFKRRGGIGDMTPSERLINMITTEQEIQLQAKN 960
QY	961	SKLKDFRYVFRGRDQLWKGPCELLWKGEAVLVKVGTDIKIIPRRKAKIIRDYGRQEM 1020
DB	961	SKLKDFRYVFRGRDQLWKGPCELLWKGEAVLVKVGTDIKIIPRRKAKIIRDYGRQEM 1020
QY	1021	DSGSHLEGAREDEGEMA 1036
DB	1021	DSGSHLEGAREDEGEMA 1036
RESULT 2		
AAY51978		
ID	AAY51978	standard; Protein; 1036 AA.
XX	AAY51978;	XX
AC	AC	XX
DT	DT	XX
DT	11-JUL-2000 (first entry)	XX
DE	HIV-2 ROD isolate polymerase DNA fragment.	XX
DE	Diagnosis; probe; lymphocyte; virus; immunodiagnostic; infection;	XX
KW	antiviral; polymerase.	XX
KW	Human immunodeficiency virus type 2.	XX
OS	US6054565-A.	XX
PN	25-APR-2000.	XX
XX	28-APR-1994; 94US-0234875.	XX
XX	11-FEB-1987; 87US-0013477.	XX
PR	03-SEP-1991; 91US-0752368.	XX
PR	20-DEC-1991; 91US-0810908.	XX
PR	03-MAR-1986; 86US-0835228.	XX
PR	06-OCT-1986; 86US-0916080.	XX
PR	21-NOV-1986; 86US-0933184.	XX
PR	16-JAN-1987; 87US-0003764.	XX
XX	(INSP) INST PASTEUR.	XX
XX	Montagnier L, Clavel F, Guyader M, Goutard D, Sonigo P, Allizon M;	XX
XX	WPI; 2000-328365/28.	XX
XX	Novel cloned nucleotide sequences homologous or identical to the	XX
PT	portion of genomic RNA of HIV-2 viruses useful as probes and in	XX
PT	diagnostic tests to diagnose HIV-2 infection	XX
XX	Claim 1; Column 11-16; 33pp; English.	XX
PS	This invention describes a novel cloned nucleic acid (I) of a human	XX
CC	immunodeficiency virus type 2 (HIV-2). (I) is capable of being used	XX
CC	as probes in diagnostic method to obtain the immunological reagents	XX
CC	necessary to diagnose an HIV-2 infection. These sequences may be used	XX
CC	as probes in hybridization reactions with the genetic material of	XX
CC	infected patients to indicate whether the RNA of the HIV-2 virus is	XX
CC	present in these patient's lymphocytes or whether an analogous DNA is	XX
CC	present. The genetic sequence of the HIV-2 virus may be used to create	XX
CC	the polypeptides encoded by these sequences. Specifically, these	XX
CC	polypeptides may be created by expression of the cDNA obtained from	XX
CC	bacterial, yeast or animal cells. These polypeptides may be used in	XX
CC	diagnostic tests such as immunofluorescence assays, radioimmunoassays	XX
CC	(RIA) and Western Blot tests. Monoclonal antibodies to these polypeptides	XX
CC	of fragments may be created and used in immunodiagnostic tests. The	XX
CC	polypeptides of the present invention may also be used as immunogenic	XX
CC	reagents to induce protection against infection by HIV-2 viruses. The	XX
CC	polypeptides produced by recombinant-DNA techniques would function as	XX
CC	vaccine agents. The polypeptides may be used on competitive assays to	XX

CC test the ability of various antiviral agents to determined their ability
CC to prevent the virus from fixing on its target. This sequence represents
CC a HIV-2 ROD isolate protein encoded by a polymerase fragment which is
CC described in the method of the invention.
XX
SQ Sequence 1036 AA;

Query Match		99.5%;	Score 5426;	DB 21;	Length 1036;
Best Local Similarity		99.6%;	Pred. No. 0;		
Matches 1032;		Conservative 1;	Mismatches 3;	Indels 0;	Gaps 0;
Qy	1	TGFRFRGPGKEAPQLPRGPSSAGADTNSPSSSGSGTGEIYAAREKTERABRETIQG	60		
Db	1	TGFRFRGPGKEAPQLPRGPSSAGADTNSPSSSGSGTGEIYAAREKTERABRETIQG	60		
Qy	61	SDRGLTAPRAGGDTIQGATNGLAAPQFSLWKRPVVTAYIEGQPVVELLDTGADDSIVAG	120		
Db	61	SDRGLTAPRAGGDTIQGATNGLAAPQFSLWKRPVVTAYIEGQPVVELLDTGADDSIVAG	120		
Qy	121	IELGNNSPKTVGGIGGFINTKEYKNVEIEVLNKKVRATIMTGTPTNIFGRNLTALGM	180		
Db	121	IELGNNSPKTVGGIGGFINTKEYKNVEIEVLNKKVRATIMTGTPTNIFGRNLTALGM	180		
Qy	181	SLNLPVAKVEPIKIMLPKGDGPKLRWPLTKIEALKEICEKMEKEGQLEEAAPPTNPY	240		
Db	181	SLNLPVAKVEPIKIMLPKGDGPKLRWPLTKIEALKEICEKMEKEGQLEEAAPPTNPY	240		
Qy	241	NTPTFAIKKDKNWRMLIDPRELNKVTQDTEIQLGIPHPAGLAKKRITVLDVGAYF	300		
Db	241	NTPTFAIKKDKNWRMLIDPRELNKVTQDTEIQLGIPHPAGLAKKRITVLDVGAYF	300		
Qy	301	SIPLHEDFRPYATFTLPSVNNAPGKRYIYKVLPGQWKGSPAIQHTMRQVLEPFRANK	360		
Db	301	SIPLHEDFRPYATFTLPSVNNAPGKRYIYKVLPGQWKGSPAIQHTMRQVLEPFRANK	360		
Qy	361	DVIIIQWDDILIASDRTDLEHDRVQLKELLNGLGFSTPDEKFPQDPPYHWMGYELWP	420		
Db	361	DVIIIQWDDILIASDRTDLEHDRVQLKELLNGLGFSTPDEKFPQDPPYHWMGYELWP	420		
Qy	421	TKWLKQLIQLPQKEIWTNDIQKLVGLVNWAAQLYPGIKTKHLCLRLINGKMTLFEVQWT	480		
Db	421	TKWLKQLIQLPQKEIWTNDIQKLVGLVNWAAQLYPGIKTKHLCLRLINGKMTLFEVQWT	480		
Qy	481	ELAEAELEENRIILSQEGHYOEEKELEATVQKDQENQWYKIHOEEKILKVGKYAKV	540		
Db	481	ELAEAELEENRIILSQEGHYOEEKELEATVQKDQENQWYKIHOEEKILKVGKYAKV	540		
Qy	541	KNTHNGIRLLAQVVOQIGKEALVINGRIPKPHLPVEREIEWQWMDNYQVWTIPDWDV	600		
Db	541	KNTHNGIRLLAQVVOQIGKEALVINGRIPKPHLPVEREIEWQWMDNYQVWTIPDWDV	600		
Qy	601	STPPLVRLAFLNGDPIPGAETFTYDGSNRSQREGAGYVYDRGDKVKKLEQTTNQQA	660		
Db	601	STPPLVRLAFLNGDPIPGAETFTYDGSNRSQREGAGYVYDRGDKVKKLEQTTNQQA	660		
Qy	661	ELEAFAMALDSGPKVNIIVDSQVVMGISASQTESKIVNQIIEEMIKKEALYVAVWP	720		
Db	661	ELEAFAMALDSGPKVNIIVDSQVVMGISASQTESKIVNQIIEEMIKKEALYVAVWP	720		
Qy	721	AHKGIGGNQEVDLHVSQIGIRQVLFLEKIEPAQEHEHYHSNVKELSHKFGIPNLVARQIV	780		
Db	721	AHKGIGGNQEVDLHVSQIGIRQVLFLEKIEPAQEHEHYHSNVKELSHKFGIPNLVARQIV	780		
Qy	781	NSCAQCOQKEAIGHQVNAELGTWQMDCTHLEGIILIVAVHVASGFTAEAVIPQESGRQT	840		
Db	781	NSCAQCOQKEAIGHQVNAELGTWQMDCTHLEGIILIVAVHVASGFTAEAVIPQESGRQT	840		
Qy	841	ALFLLLKSLASRPITHLTDNGANFTSQEVKVMVNWIGIEQSGFYYPNPQSGVGYEAMNHH	900		
Db	841	ALFLLLKSLASRPITHLTDNGANFTSQEVKVMVNWIGIEQSGFYYPNPQSGVGYEAMNHH	900		
Qy	901	LKNOISIRIQANTIETIVLMAIHCNMFKRRGGIDMTSPSRLLINMITTEQEIQFLQAKN	960		

Db	901	LKNEISIRBQANTIETIVLMAIHCNMFKRRGEIGDMTPSERLINMITTQEIQFLQAKN	960		
Qy	961	SKLKDFRVYFREGDQLWKGPGLLWKGEAVLVKVGTDIKIIPRRKAKIIRDYGGROEM	1020		
Db	961	SKLKDFRVYFREGDQLWKGPGLLWKGEAVLVKVGTDIKIIPRRKAKIIRDYGGROEM	1020		
Qy	1021	DSGSHLEGAREDEGEMA 1036			
Db	1021	DSGSHLEGAREDEGEMA 1036			
RESULT 3					
AAP81773					
ID	AAP81773 standard; protein; 1027 AA.				
XX	AAP81773;				
XX					
DT	25-MAR-2003 (updated)				
DT	15-NOV-1990 (first entry)				
XX					
DE	Sequence encoded by open reading frame of cDNA corresponding to				
DE	HIV-2 ROD genome.				
XX					
KW	LAV-II ROD; AIDS; immunogen; antigen; vaccine; diagnostic.				
XX					
OS	Human immunodeficiency virus ROD.				
XX					
PN	WO8805440-A.				
PN					
PD	28-JUL-1988.				
XX					
PF	15-JAN-1988; 88WO-EP00025.				
XX					
PR	16-JAN-1987; 87US-0003764.				
PR	11-FEB-1987; 87PR-0001739.				
PR	15-APR-1987; 87PR-0005398.				
XX					
PA	(INSP) INST PASTEUR.				
PA	(ALIZ/) ALIZON M.				
PA	(CNRS) CENT NAT RECH SCI.				
XX					
PI	Alizon M, Montagnier L, Guetard D, Clavfl F, Sonigo P, Guyader M;				
PI	Tiollais P, Chakrabarti L, Desrosiers R;				
XX					
DR	WPI: 1988-220290/31.				
DR	N-PSDB; AAN80859.				
XX					
PT	New peptides with immunological properties of HIV-2 envelope protein -				
PT	having the structure of simian immune deficiency virus proteins,				
PT	useful in diagnosis and of vaccine components				
XX					
PS	Disclosure; Fig 1A; 86pp; French.				
XX					
CC	The SQ in AAN80859 was deposited on 21/2/86 at the CNCM under number				
CC	I-522, reference name LAV-II ROD. It is the cDNA to HIV-2 ROD genomic				
CC	RNA. The SQ was compared with the SQ of the genome of SIV (Mac) (AAN80860)				
CC	to identify common regions.				
CC	(Updated on 25-MAR-2003 to correct PF field.)				
CC	(Updated on 25-MAR-2003 to correct PR field.)				
CC	(Updated on 25-MAR-2003 to correct PA field.)				
XX					
SQ	Sequence 1027 AA;				
Query Match					
Best Local Similarity 98.7%; Score 5378.5; DB 9; Length 1027;					
Matches 1023; Conservative 2; Mismatches 2; Indels 9; Gaps 1;					
Qy	1	TGFRFTGPGKEAPQLPRGPSSAGADTNSPSSSGSGTGEIYAAREKTERABRETIQG	60		
Db	1	TGFRFTGPGKEAPQLPRGPSSAGADTNSPSSSGSGTGEIYAAREKTERABRETIQG	60		
Qy	61	SDRGLTAPRAGGDTIQGATNRLGAAPQFSLWKRPVVTAYIEGQPVVELLDTGADDSIVAG	120		


```
Db 241 ELNKVTDQTEIQLGIPHPAGLAKRRRTITVLDVGDAFYFSIPLHEDFRPYTAFTLPSVNA 300
Qy 323 EPGRRYIKVLPQGWKSPALFQHTMRQVLEPPFRKANKDVIIIOYMODILLASDRTLEH 382
Db 301 EPGRRYIKVLPQGWKSPALFQHTMRQVLEPPFRKANKDVIIIOYMODILLASDRTLEH 360
Qy 383 DRVVLQKELLNGLGFTPDDEKFKQDPPYHWMGYELWPTKWKLOKIQLPQKEIWTVDIQ 442
Db 361 DRVVLQKELLNGLGFTPDDEKFKQDPPYHWMGYELWPTKWKLOKIQLPQKEIWTVDIQ 420
Qy 443 KLGVGLNAAQLYPGIKTKHLCLRLIRKMWLTTEEVQWTELAELNRIILSQEGHY 502
Db 421 KLGVGLNAAQLYPGIKTKHLCLRLIRKMWLTTEEVQWTELAELNRIILSQEGHY 480
Qy 503 YOEEKELEATVOKDOENQWYKIHOEEKILKVGYAKVKNHTNGIRLLAOVQKIGEA 562
Db 481 YOEEKELEATVOKDOENQWYKIHOEEKILKVGYAKVKNHTNGIRLLAOVQKIGEA 540
Qy 563 LVIWGRIPKPHLPVEREIQWMDNYQVWTWIPDWFVSTPPLVRLAFNLVGDPIPAET 622
Db 541 LVIWGRIPKPHLPVEREIQWMDNYQVWTWIPDWFVSTPPLVRLAFNLVGDPIPAET 600
Qy 623 FYTDCSNROSKEGKAGYVTDGRKDKVKLEQTTNOQAELEAFAMALTDSPKVIIVDS 682
Db 601 FYTDCSNROSKEGKAGYVTDGRKDKVKLEQTTNOQAELEAFAMALTDSPKVIIVDS 660
Qy 683 QYVNGISASOPTESKIVNOIIEEMIKKEAIYVAVWPAHKIGGNOEVDHLVSGIRQV 742
Db 661 QYVNGISASOPTESKIVNOIIEEMIKKEAIYVAVWPAHKIGGNOEVDHLVSGIRQV 720
Qy 743 LFLKIEPAQEHEKYSNVKELSHKFGIPNLVARIQVNSCAQCOQKGEALHGVNAELG 802
Db 721 LFLKIEPAQEHEKYSNVKELSHKFGIPNLVARIQVNSCAQCOQKGEALHGVNAELG 780
Qy 803 TWQMDCTHLEGGKIIIVAVHVASGFIEAIVIPQESGROTALFLLKLASRWPTTHLTDNGA 862
Db 781 TWQMDCTHLEGGKIIIVAVHVASGFIEAIVIPQESGROTALFLLKLASRWPTTHLTDNGA 840
Qy 863 NTSQEVKMAWNIIGIEQSGVPYNPOSQGVVEAMNHLKNOISRIREQANTIETVIMA 922
Db 841 NTSQEVKMAWNIIGIEQSGVPYNPOSQGVVEAMNHLKNOISRIREQANTIETVIMA 900
Qy 923 IHCNFKRRGGIGDMTPSERLINMITTEQETQFLOAKNSKLKDPVYFREGRODLWKPG 982
Db 901 IHCNFKRRGGIGDMTPSERLINMITTEQETQFLOAKNSKLKDPVYFREGRODLWKPG 960
Qy 983 ELLWKGGAVLVKGTDIKIIPRRKAKIIRDYGGREQMDSGSHLEGAREDEMA 1036
Db 961 ELLWKGGAVLVKGTDIKIIPRRKAKIIRDYGGREQMDSGSHLEGAREDEMA 1014

RESULT 5
AAP93285
ID AAP93285 standard; protein: 3080 AA.
XX
AC AAP93285;
XX
DT 25-MAR-2003 (updated)
DT 17-DEC-2001 (updated)
DT 06-APR-1990 (first entry)
XX
DE Sequence of clone HIV-2 SBL/ISY.
XX
KW HIV-2; proviral clone HIV-2 SBL/ISY;
XX
OS Human immunodeficiency virus 2.
XX
PN USN7331212-N.
XX
PD 29-AUG-1989.
XX
PF 31-MAR-1989; 89US-0331212.
XX
```

```
PR 30-MAR-1989; 89US-0330446.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICE.
PA (USDC ) US SEC OF COMMERCE.
XX
PI Franchini G, Wongstaal F, Gallo R;
XX
XX WPI; 1989-339698/46.
DR N-PSDB; AAN92119.
DR
XX
PT Complete human immunodeficiency type 2 proviral clone - used to generate
XX animal model for function studies of HIV genes in vivo.
PS Disclosure; Fig. 5; 43pp; English.
XX
XX The protein is encoded by the third reading frame of HIV-2 SBL/ISY, a
CC proviral clone of HIV-2.
CC (Note: Revised entry submitted to correct the patent number format of
CC US Government-owned NTIS applications to prevent clashes with ongoing US
CC granted patent numbers. For further information please visit the Derwent
CC web site at www.derwent.com/dwpi/updates/ntis_us.html.)
CC (Updated on 25-MAR-2003 to correct PR field.)
CC (Updated on 25-MAR-2003 to correct PA field.)
CC (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ Sequence 3080 AA;
Query Match 93.1%; Score 5075.5; DB 10; Length 3080;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 950; Conservative 49; Mismatches 36; Indels 1; Gaps 1;
Qy 1 TGRFFRTCPGLGKEAPQLPRGPSSAGADTNSPSSSGSTGEIYAAREKTERAERETIQ 60
Db 584 TGWFFRAWTMKEAPQLPRGPFGAGANTNSPSSSGSTGEVHAAREKTERAETKTQR 643
Qy 61 SDRGLTAPRAGDRTIQGATNRGLAAPQFSLWKRVPVWTAYIEGQPEVLLDTCADDISVAG 120
Db 644 SDRGLAASRRARDTQ-RDDRGLAAPQFSLWKRVPVWTAYIEDQPEVLLDTCADDISVAG 702
Qy 121 IELNNYSPKIVGGIGGFINTKEYKNVEIEVLNKKVRATIMTGTPTINIFGRNLTALGM 180
Db 703 IELGNSYSPKIVGGIGGFINTKEYKDVEIRVLNKKVRATIMTGTPTINIFGRNLTALGM 762
Qy 181 SLNLPVAKVEPIKTLMPGKDGPKLRWPLTKETALKEICEKMEREGQLEEAAPPNTPY 240
Db 763 SLNLPVAKIEPVKTLMPGKDGPKQRWPLTKETALKEICEKMEREGQLEEAAPPNTPY 822
Qy 241 NTPPTFAIKKKDKNKRMLIDFRELNKVTDQFTEIQLGIPHPAGLAKRRRTITVLDVGDAFY 300
Db 823 NTPPTFAIKKKDKNKRMLIDFRELNKVTDQFTEIQLGIPHPAGLAKRRRTITVLDVGDAFY 882
Qy 301 SIPLHEDFRPYTAFTLPSVNNAPGKRYIKVLPQGWKSPALFQHTMRQVLEPPFRKANK 360
Db 883 SIPLYEDFRQYTAFTLPSVNNAPGKRYIKVLPQGWKSPALFQHTMRQVLEPPFRKANK 942
Qy 361 DVIIIOYMDILLIASDRTLEHDKVILQKELLNGLGFTPDDEKFKQDPPYHWMGYELWLP 420
Db 943 DVIIIOYMDILLIASDRTLEHDKVILQKELLNGLGFTPDDEKFKQDPPYHWMGYELWLP 1002
Qy 421 TKWKLQIKQLPQKEIWTVDIQLKLVGLNAAQLYPGIKTKHLCLRLIRKMWLTTEEVQWT 480
Db 1003 TKWKLQIKQLPQKEIWTVDIQLKLVGLNAAQLYPGIKTKHLCLRLIRKMWLTTEEVQWT 1062
Qy 481 ELAAELEENRIILSQEGHYHOEEKELEATVOKDOENQWYKIHOEEKILKVGKAKV 540
Db 1063 ELAAELEENRIILSQEGHYHOEEKELEATVOKDOENQWYKIHOEEKILKVGKAKV 1122
Qy 541 KNHTNGIRLLAOVQKIGKALYVIRPKFHLPVVEREIQWMDNYQVWTWIPDMDFFV 600
Db 1123 KNHTNGIRLLAOVQKIGKALYVIRPKFHLPVVEREIQWMDNYQVWTWIPDMDFFV 1182
Qy 601 STPLPLVRLAFNLVGDPIPGAETFTDGCNSRQSKGKAGYVTDGRKDKVKLEQTTNOQAE 660
Db 1182 STPLPLVRLAFNLVGDPIPGAETFTDGCNSRQSKGKAGYVTDGRKDKVKLEQTTNOQAE 660
```

Db 1183 STPLVRLAFNLVKDPIPGAETFTYTDGSCNROSKEGAGYITDRGKDKVRILEQTNTNOQA 1242
Qy 661 ELEAFAMALTDSPGKNIIVDSQYVMGISASQPTSESKIVNQIIEEMIKKEAIYVAVWP 720
Db 1243 ELEAFAMAYTDSGPKNIIVDSQYVMGIVTGQPAESERIVNKKIIEEMIKKEAIYVAVWP 1302
Qy 721 AHKGIGNQEVHLVSGIRQVLFLEKIPQAEEHEKYHNSVKELSHKFGIPNLVARQIV 780
Db 1303 AHKGIGNQEVHLVSGIRQVLFLEKIPQAEEHEKYHNSVKELSHKFGIPNLVARQIV 1362
Qy 781 NSCAOCCQKGEATHGQVNAELGTWQMDCTHLEKIIIVAVHVASGFIEAEVPOESGROT 840
Db 1363 NTCACQCCQKGEATHGQVNAELGTWQMDCTHLEKIIIVAVHVASGFIEAEVPOESGROT 1422
Qy 841 ALFLLKLASRPWITHLHTDNGANFTSQEYKVMVAMWIGIEQSFQVYPNPOSQGVVEAMNHH 900
Db 1423 ALFLLKLASRPWITHLHTDNGANFTSQEYKVMVAMWIGIEQSFQVYPNPOSQGVVEAMNHH 1482
Qy 901 LKNQISRIREQANTITIVLMAIHCNMFRRRGIGDMTPSERLINMITTEQEIQFLQAKN 960
Db 1483 LKNQIERIREQANTMETIVLMAVHCNMFRRRGIGDMTPSERLINMITTEQEIQFLQAKN 1542
Qy 961 SKLKDPFRVYFREGRODLWGPGLLWKGEGAVLVKVGTDIKIIPRKAKIIRDYGGROEM 1020
Db 1543 SKLKDPFRVYFREGRODLWGPGLLWKGEGAVLVKVGTDIKIIPRKAKIIRDYGGROEM 1602
Qy 1021 DSGSHLEGAREDEGEMA 1036
Db 1603 DSGSHLEGAREDEGEMA 1618

RESULT 6

AAAP81771
ID RAA81771 standard; protein; 3210 AA.
XX
AC AAP81771;
XX
XX 25-MAR-2003 (updated)
DT 07-NOV-1990 (first entry)
XX
DE Deduced sequence encoded by bottom reading frame of cDNA clone
DE HIV-2 SBL/ISY of HIV related retrovirus strain.
XX
KW HIV vaccine; HIV strain SBL-6669-85.
XX
OS Human immunodeficiency virus.
XX
PN W08080449-A.
XX
PD 03-NOV-1988.
XX
PF 28-APR-1988; 88WO-SE00218.
XX
PR 28-APR-1987; 87SE-0001765.
XX
PA (SBL-2) SBL STATENS BAKTERI.
PA (STAT-2) STATENS BAKTERIOLOGISKA LAB.
XX
XX Albert J, Biberfeld G, Fenyo EM, Norrby E;
XX
DR WPI; 1988-322769/45.
DR N-PSDB; AAN80890.
XX
PT HIV related human retro-virus strain -
PT used for obtaining antigens for assays and vaccines and for
PT prodn. of antibodies for assays
XX
PS Claim 9; Fig 4; 28pp; English.
XX
CC Synthetically produced proteins and peptides, characterised in that the
CC AA sequence is derived from the primary nucleotide sequence of
CC HIV-2 SBL/ISY or a part thereof, or a degenerate thereof are claimed.
CC HIV-2 SBL/ISY represents the complete genome of the virus SBL-6669

CC (=SBL-6669-85). The proviral DNA was obtd. from a genomic library
CC constructed from DNA of HUT-78 cells infected with SBL-6669-85 using
CC the lambda-phage vector EMBL-3. SBL-6669-85 was isolated from lymphocytes
CC of a West African woman. Protection is requested for the entire genome
CC disclosed in n80890 and for parts thereof, and corresp. to various genes
CC such as the gag gene (corresp. to nucleotides 547 to 2106), the pol gene
CC (nucleotides 1827-4931) and the env gene (nucleotides 6144 to 8682), the
CC corresp. AA sequences and parts thereof and various products derived
CC therefrom, or use thereof, such as clones prepd. by recombinant vector
CC method, HIV test devices and methods. X corresponds to the translation of
CC a stop codon.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX

SQ Sequence 3210 AA;

Query Match 92.8%; Score 5059; DB 9; Length 3210;
Best Local Similarity 91.6%; Pred. No. 0;
Matches 949; Conservative 49; Mismatches 36; Indels 2; Gaps 2;

Qy 1 TGRFRTGPLGKEAPQLPRGPSSAGADTNSTPSGSSSGSTGEBIYAAREKTERAETIQG 60
Db 609 TGMFFRAMTGMKEAPQLPRGPKEFAGANTNSTPNSGSSGPTGVBHAAREKTERAETKTQR 668
Qy 61 SDRGLTAPAGDGTTOGATNRCGLAAPQSLKRPVVTAYIEQGPVEVLLDTGADDSIVAG 120
Db 669 SDRGLAARARDDTTQ-RDDRGLAAPQSLKRPVVTAYIEQGPVEVLLDTGADDSIVAG 727
Qy 121 IELGNNYSPKIVGGIGGFGINTKEYKNVEIEVLNKKVRATIMTGDPTINIFGNILTALGM 180
Db 728 IELGNNYSPKIVGGIGGFGINTKEYKNVEIEVLNKKVRATIMTGDPTINIFGNILTALGM 787
Qy 181 SLNLPVAKVEPIKIMLPKPGDKPKLRQWPLTKETKTEALKEICEKMEKEQLEAPPTNPY 240
Db 788 SLNLPVAKIEPVKVTLPKPGDKPKLRQWPLTKETKTEALKEICEKMEKEQLEAPPTNPY 847
Qy 241 NTPTEAIIKKKKKWRMLIDFRELNKVTQDFTFVGLGPHPAAGLAKKRITVLDVGDYAF 300
Db 848 NTPTEAIIKKKKKWRMLIDFRELNKVTQDFTFVGLGPHPAAGLAKKRITVLDVGDYAF 907
Qy 301 SIPLHEDFRPYTAFTLPSVNAEPCKRYIKVLPQWKGSPAIQFQHTWRQVLEPRKANK 360
Db 908 SIPLYEDFRPYTAFTLPSVNAEPCKRYIKVLPQWKGSPAIQFQHTWRQVLEPRKANK 967
Qy 361 DVIIQYMDILLIASDRDLDHDRVVLQKELLNGLSGFSTPDEKFKQDPPYHMMGYELWP 420
Db 968 DVIIQYMDILLIASDRDLDHDRVVLQKELLNGLSGFSTPDEKFKQDPPYHMMGYELWP 1027
Qy 421 TKWKLOKIQLOKEITWTVNDIQKLVGLNWAQLYPGIKTKHLCLIRGKMTLITEVQMT 480
Db 1028 TKWKLOKIQLOKEITWTVNDIQKLVGLNWAQLYPGIKTKHLCLIRGKMTLITEVQMT 1087
Qy 481 ELAEAELEENRIILSQEGBHYQBEKELEATVQKQDNQWYKTHQEEKILKVGKYAKV 540
Db 1088 ELAEAELEENRIILSQEGBHYQBEKELEATVQKQDNQWYKTHQEEKILKVGKYAKV 1147
Qy 541 KNTHTNGIRLLAQVYVQKIGKEALVIGWRIPKFLPVEREIBEQWMDNWTWQVTPDWDV 600
Db 1148 KNTHTNGIRLLAQVYVQKIGKEALVIGWRIPKFLPVEREIBEQWMDNWTWQVTPDWDV 1207
Qy 601 STPLVRLAFNLVGDPIPGAETFTYTDGSCNROSKEGAGYITDRGKDKVRILEQTNTNOQA 660
Db 1208 STPLVRLAFNLVGDPIPGAETFTYTDGSCNROSKEGAGYITDRGKDKVRILEQTNTNOQA 1267
Qy 661 ELEAFAMALTDSPGKNIIVDSQYVMGISASQPTSESKIVNQIIEEMIKKEAIYVAVWP 720
Db 1268 ELEAFAMALTDSPGKNIIVDSQYVMGISASQPTSESKIVNQIIEEMIKKEAIYVAVWP 1327
Qy 721 AHKGIGNQEVHLVSGIRQVLFLEKIPQAEEHEKYHNSVKELSHKFGIPNLVARQIV 780
Db 1328 AHKGIGNQEVHLVSGIRQVLFLEKIPQAEEHEKYHNSVKELSHKFGIPNLVARQIV 1387
Qy 781 NSCAOCCQKGEATHGQVNAELGTWQMDCTHLEKIIIVAVHVASGFIEAEVPOESGROT 840
XX :|||||

Db	1388	NTCAOQCKGEALHGVNAELGTWQMDCTHLEGGKIIIVAVHVASGFTAEAVIPQESGRQT	1447
QY	841	ALFLLKASRWPIITHLTDNGANFTSQEYKVMYANWIGIEQSGFVYPNPQSGVVEAMNHH	900
Db	1448	ALFLLKASRWPIITHLTDNGANFTSQEYKVMYANWIGIEQSGFVYPNPQSGVVEAMNHH	1507
QY	901	LKNQISRIREQANTTETIVLMAHCHMNFKRRGGIGDWTTPSERLINMTTEQEIFLOAKN	960
Db	1508	LKNQISRIREQANTTETIVLMAHCHMNFKRRGGIGDWTTPSERLINMTTEQEIFLOAKN	1567
QY	961	SKLKDFRVYFRGRDQWKPGCELLWKGGAVLVKVGTDIKIIPRRKAKIIRDYGGROEM	1020
Db	1568	SKLKNFRVYFRGRDQWKPGCELLWKGGAVLVKVGTDIKIIPRRKAKIIRDYGGROEM	1627
QY	1021	DSGSHLEGAREDEGEMA 1036	
Db	1628	DSGSHLEGARE-DEGEMA 1642	
RESULT 7			
ID	AAW13055		
AC	AAW13055 standard; Protein; 1055 AA.		
AC	AAW13055;		
XX			
DT	25-MAR-2003 (updated)		
DT	15-OCT-1997 (first entry)		
DE	HIV-2 provirus-encoded pol protein.		
XX			
KW	HIV-2; HIV-2KR; provirus; gene therapy; vector: vaccine; AIDS;		
KW	pol protein.		
XX			
OS	Human immunodeficiency virus type 2.		
XX			
PN	WQ9705242-AL.		
PD	13-FEB-1997.		
XX			
PF	09-JUL-1996; 96WO-US11445.		
PR	07-JUN-1996; 96US-0659251.		
PR	26-JUL-1995; 95US-0001441.		
XX			
PA	(REGC) UNIV CALIFORNIA.		
XX			
PI	Kraus G, Wongstaal F, Talbott R, Poeschla E;		
XX			
DR	WPI; 1997-145681/13.		
DR	N-PSDB; AAT61085.		
XX			
PT	Human immunodeficiency virus type II proviral clone - useful to		
PT	develop prods. for study, diagnosis, prevention and treatment of HIV		
PT	infections		
XX			
PS	Disclosure; Page 89-92; 128pp; English.		
XX			
CC	Polypeptides (AAW13052-59) respectively comprise the env, gag, nef,		
CC	pol, rev, tat, vif and vpr polypeptides encoded by HIV-2KR (AAT61085),		
CC	an HIV-2 provirus that comprises a full-length HIV-2 genome.		
CC	HIV-2KR was obt'd. from the genomic DNA of Molt 4 Clone 8 lymphoblast		
CC	cells infected with HIV-2P12, an HIV-2 strain isolated from the		
CC	peripheral blood mononuclear cells of an HIV-2 seropositive AIDS		
CC	patient. HIV-2KR polypeptides can be expressed in host cells and		
CC	used as diagnostic reagents and as immunogens for the production of		
CC	diagnostic antibodies.		
CC	(updated on 25-MAR-2003 to correct PI field.)		
XX			
XX	Sequence 1055 AA;		
Query Watch	92.7%; Score 5055.5; DB 18; Length 1055;		
Best Local Similarity	91.5%; Pred. No. 0;		
Matches 948; Conservative	44; Mismatches 43; Indels 1; Gaps 1;		

RESULT 8	
AA04025	
ID	AA04025 standard; protein; 1035 AA.
XX	
AC	AA04025;
XX	
DT	25-MAR-2003 (updated)
DT	29-MAY-1989 (first entry)
XX	
DE	Pol gene product of cDNA to HIV-2 RNA.
XX	
KW	HIV; AIDS; Vaccine; pUC-HIV-2(GH-1).
XX	
OS	HIV-2.
XX	
PN	JP01289486-A.
XX	
PD	21-NOV-1989.
XX	
PF	16-MAY-1988; 88JP-0119024.
XX	
PR	16-MAY-1988; 88JP-0119024.
XX	
PA	(TOFU) TOA NENRYO KOGYO KK.
PA	(FJRE) FUJI REBIO KK.
XX	
DR	WPI; 1990-005177/01.
DR	P-PSDB; AA040424, AA040425, AA040426, AA040427, AA040428, AA040429.
XX	
PT	DNA indicating complement to RNA gene -
PT	of Human Immunodeficiency Virus type 2 used for new vaccine or
PT	diagnostic for AIDS virus.
XX	
PS	Claim 2; Fig.4; 12pp; Japanese.
XX	
CC	cDNA to novel HIV-2 (GH-1) has been integrated into plasmid
CC	pUC HIV-2(GH-1). Useful for diagnosis and vaccination against the
CC	virus. Described here is the Pol gene product of the cDNA.
CC	See also AA002830.
CC	(Updated on 25-MAR-2003 to correct PA field.)
XX	
SQ	Sequence 1035 AA;
	Query Match 92.1%; Score 5023.5; DB 11; Length 1035;
	Best Local Similarity 91.0%; Pred. No. 0;
	Matches 944; Conservative 52; Mismatches 38; Indels 3; Gaps 3;
Qy	1 TGRFFRTGPGKEAPOLPRGPSSGADTNTSPSSSGSTGETIYAAREKTERAERTIQG 60
Db	1 TGRFFRDGSGKEAPOLPRGPSSGADTNTSPSSSGSTGKIYAAGERAEAGETIQR 60
Qy	61 SDRGLTAPRAGDGTIOGATNRGLAAPQFSLWKRPPVTVATIEGQPVVLLDTGADDSIVAG 120
Db	61 GDGRLTAPRAGKSTSG-DRGLAAPQFSLWKRPPVTVATIEQPVVLLDTGADDSIVAG 119
Qy	121 IELGNNYSPKIVGGIGFINTKEYNVEVLNKKVRATIMTGDTPINIFGRNILTALGM 180
Db	120 IQLGDNVYKIVGGIGFINTKEIKNIEKVLNKRVRATIMTGDTPINIFGRNILTALGM 179
Qy	181 SLNLPVAKVEPKIMLKPGKDGPKLRQWPLTKIEALKEICEKMEKEGQLEAPPTNY 240
Db	180 SLNLPVAKVEPKIMLKPGKDGPKLRQWPLTKIEALKEICEKMEKEGQLEAPPTNY 239
Qy	241 NTPTEFAIKKKDKNKRMLIDFELNKNVTDFTIEQLGIPHPAGLAKKRITVLDVGDYF 300
Db	240 NTPTEFAIKKKDKNKRMLIDFELNKNVTDFTIEQLGIPHPAGLAKKRITVLDVGDYF 299
Qy	301 SIPLHEDFRPYTAFTLPSVNNAPPGKRYIKVLPQGWKGSIPAIFQHTMRQVLEPFRKANK 360
Db	300 SIPLHEDFRPYTAFTLPSVNNAPPGKRYIKVLPQGWKGSIPAIFQHTMRQVLEPFRKANK 359
Qy	361 DVIIIQYMDILLIASDRTGLEHDKVYVQLKELLNGLGFSTPDEKFKQDPLQMMGYELWP 420

Db	360 DVIIIQYMDILLIASDRTGLEHDKVYVQLKELLNGLGFSTPDEKFKQDPLQMMGYELWP 419
Qy	421 TWKQLQKIQLOPOKEIWTWVNDIOKLVGLNWAOLYPGIKTKHLCLRLIRCKMTLITEVQWT 480
Db	420 TWKQLQKIQLOPOKEIWTWVNDIOKLVGLNWAOLYPGIKTKHLCLRLIRCKMTLITEVQWT 479
Qy	481 ELAEAELEENRITLSQEQEGHYQEEKELEATVQKDOENQWYKIHQEEKILKVGKYAKV 540
Db	480 ELAEAELEENKILSQEQEGSYQEEKELEATIQKQDNQWYKIHQEEKILKVGKYAKI 539
Qy	541 KNTHTNGIRLLAQVQKIGKEALVINGRIPKPHLPVEREIEWQWMDNYQVWTIPDQWF 600
Db	540 KNTHTNGVRLLAQVQKIGKEALVINGRIPKPHLPVERETWEQWMDNYQVWTIPDQWF 599
Qy	601 STPPLVRLAFNLVGDPIPGAETFTYDGSNRSKSGKAGYVTDGRGDKVKKLEOTTNQQA 660
Db	600 STPPLVRLTFNLVGDPIPGAETFTYDGSNRSKSGKAGYVTDGRGDKVRLERTNQQA 659
Qy	661 ELEAFAMALTDGPKVNIIVDSQYVMGISASQPTSESKIVNQIIEEMIKKEAIVYAWVP 720
Db	660 ELEAFAMALTDGPKVNIIVDSQYVMGIVVGQPTSESRIVNQIIEDMIKKEAVYAWVP 719
Qy	721 AHKGIGGNOEVDHLVSGQIRQVLFLEKIEPAQEEHEKYHSNVKELSHKEGINLVARQIV 780
Db	720 AHKGIGGNOEVDHLVSGQIRQVLFLEKIEPAQEEHEKYHSNMKELTHKFGIPQLVARQIV 779
Qy	781 NSCAOQOQGEATHGOVNAELGTWQMDCTHLEGIITIVAVHVASGFIEAEVPOESGROT 840
Db	780 NTCACQOQGEATHGOVNAELGTWQMDCTHLEGIITIVAVHVASGFIEAEVPOESGROT 839
Qy	841 ALFLKLASRPITHLHTDNGANFTSQEVKVMVWVIGIEQSEFGVPYNPQSQGVVEAMNHH 900
Db	840 ALFLKLASRPITHLHTDNGSFTSQEVKVMVWVIGIEQSEFGVPYNPQSQGVVEAMNHH 899
Qy	901 LKNQISRIREQANTTETIVLMAIHCNFKRRGIGDMTSPERLINMITTEQEIQFLOAKN 960
Db	900 LKNQISRIREQANTTETIVLMAVHCNFKRRGIGDMTPAERLINMITTEQEIQFLOAKN 959
Qy	961 SKLKDFRVYVREGDQLWKPGELLMKGGAVLVKVTGDIKIIPRKAIIIRDYGGROEM 1020
Db	960 SNFKNFQVYVREGDQLWKPGELLMKGGAVLVKVGADIKVIPRKAIIIRDYGGRQEL 1019
Qy	1021 DSGSHLEGAR-EDGEMA 1036
Db	1020 DS-SHLEGAREEDGEVA 1035
RESULT 9	
AAW89314	
ID	AAW89314 standard; Protein; 1060 AA.
XX	
AC	AAW89314;
XX	
DT	01-JUN-1999 (first entry)
XX	
DE	SIVmac239 genome pol protein sequence.
XX	
KW	Antigenic composition; primate; lentivirus; nef gene; vaccine;
XX	Infection; AIDS; SIVmac239; pol protein.
OS	Simian immunodeficiency virus.
XX	
PN	US5851813-A.
PD	22-DEC-1998.
XX	
PF	27-JAN-1994; 94US-0188583.
XX	
PR	27-JAN-1994; 94US-0188583.
PR	12-JUL-1990; 90US-0551945.
PR	09-JUL-1991; 91US-0727494.
XX	

CC	(Updated on 25-MAR-2003 to correct PF field.)										
CC	(Updated on 25-MAR-2003 to correct PR field.)										
CC	(Updated on 25-MAR-2003 to correct PA field.)										
XX											
SQ	Sequence	1056 AA;									
<hr/>											
QY	1	TGFRFTGPGKEAPQLPRGPSSAGADTNSTPSGSSSGSTGEIYAAREKTERAERE	TIQG	60							
DB	39	TGFFRPWPLGKEAPQFPHGSSASGADANCSPRTSCGSAKELHALGQAAERKQRE	ALQG	98							
QY	61	SDRGLTAPRAGGDTIQGATNRGLAAPQFSLWKRPPVTAYIEGQPVVLLDTGADDS	IVAG	120							
DB	99	GD-----RGFAAPQFSLWRRPVVTAHIEGQPVVLLDTGADDSIVTG	140								
QY	121	IELGNNYSPKIVGIGGFINTREYKNVEIEVLNKKVRATIMTGDTPINIFGRNLTAL	GM	180							
DB	141	IELGPHYTPKIVGIGGFINTREYKNVEIEVLGKRIGTIMGDTPTINIFGRNLTAL	GM	200							
QY	181	SLNLPVAKVEPIKMLKPKGDKPLKQWPLTKKEIKALKEICEKMEKGOLEEAPT	NPY	240							
DB	201	SLNLPVAKVEPIKMLKPKGDKPLKQWPLTKKEIKALKEICEKMEKGOLEEAPT	NPY	260							
QY	241	NTPTFAIKKKDKNRWLIDFRELNVKTDFTFQIOLGIPHPAGLAKKRITVLD	VCAYF	300							
DB	261	NTPTFAIKKKDKNRWLIDFRELNVKTDFTFQIOLGIPHPAGLAKKRITVLD	VCAYF	320							
QY	301	SIPLDEHFRPYTAFTLPSVNNAPGKRYIYKVLPOGWKSPALFOHTMRQVLEP	FRKAN	360							
DB	321	SIPLDEHFRPYTAFTLPSVNNAPGKRYIYKVLPOGWKSPALFOHTMRQVLEP	FRKAN	380							
QY	361	DVIIQYMDILIASDRTDLEHVRVVLQKELLNGLFGSTPEKFKODPPYHMGVEL	WLP	420							
DB	381	DVTLVQYMDILIASDRTDLEHVRVVLQKELLNGLFGSTPEKFKODPPYHMGVEL	WLP	440							
QY	421	TKWKLQIOLPOKEIWTVDIOKLVLGNWAAQLYPGIKTKLRLIRKMTLTERV	QWT	480							
DB	441	TKWKLQIOLPOKEIWTVDIOKLVLGNWAAQLYPGIKTKLRLIRKMTLTERV	QWT	500							
QY	481	ELAEAELEENRILTSQEGHYQEKELEATVQKDQENQWYKIHOEEKILKVGK	YAKV	540							
DB	501	ELAEAELEENRILTSQEGHYQEKELEATVQKDQENQWYKIHOEEKILKVGK	YAKV	560							
QY	541	KNTHNGIRLLAQQVQKIGKEALVWGRIPKPHLPVEREIEGWQWNYQVWIPW	DFV	600							
DB	561	KNTHNGIRLLAQQVQKIGKEALVWGRIPKPHLPVEREIEGWQWNYQVWIPW	DFV	620							
QY	601	STPPLVRLAFNLVDPGAEFTYTDGSCNRQSKGAGYVTDGKDKVKKLEQTN	QQA	660							
DB	621	STPPLVRLAFNLVDPGAEFTYTDGSCNRQSKGAGYVTDGKDKVKKLEQTN	QQA	680							
QY	661	ELAEAFALMTDGPKNIIIVDSQVVMGIIISASOPTESKIVNQIIEEMIKKE	AIYVAV	720							
DB	681	ELAEAFALMTDGPKNIIIVDSQVVMGIIISASOPTESKIVNQIIEEMIKKE	AIYVAV	740							
QY	721	AHKGIGGNQVHLVSGQIRQVLFLEKIEPAQEHEKYHSNVKELSHKFGINL	VARQIV	780							
DB	741	AHKGIGGNQVHLVSGQIRQVLFLEKIEPAQEHEKYHSNVKELSHKFGINL	VARQIV	800							
QY	781	NSCAOQOQGEATHGVNAELGTWQMDCTHLEKIIIVAVHVASGFIEAEV	IPQESGR	840							
DB	801	NSCAOQOQGEATHGVNAELGTWQMDCTHLEKIIIVAVHVASGFIEAEV	IPQESGR	860							
QY	841	ALFLLKLASRWPTLHLDGANGFTSQEYKVMVAVWITGIEQSFQVYPN	QSQGV	900							
DB	861	ALFLLKLASRWPTLHLDGANGFTSQEYKVMVAVWITGIEQSFQVYPN	QSQGV	920							
QY	901	LKNQISRIEAOATIEITVLMATHCMNFRGGIGDMTSPERLINMITTEQ	EOITQ	960							
DB	921	LKNQISRIEAOATIEITVLMATHCMNFRGGIGDMTSPERLINMITTEQ	EOITQ	980							

QY 361 DVIIIOYMDILLASDRDLEHDRVVLQKLELNGLFSTPDEKFOKDPYHMMGYELWP 420
 DB 381 DVTLVQYMDILLASDRDLEHDRVVLQKLELNGLFSTPDEKFOKDPYHMMGYELWP 440
 QY 421 TKWKLQIQLPQKEITVNDIOKLGVLNWAAQYFGIKTKHLCLIRIGKMTLTERVQWT 480
 DB 441 IKWKLQIQLPQKEITVNDIOKLGVLNWAAQYFGIKTKHLCLIRIGKMTLTERVQWT 500
 QY 481 ELAEAELEENRIILSFOEGHYQEEKELEATVQKDOENQWTYKIHQEEKILKVGKAYK 540
 DB 501 EMAEAEYENKIILSFOEGHYQEEKELEATVQKDOENQWTYKIHQEEKILKVGKAYK 560
 QY 541 KNTHTNGIRLLAQQVQKIGKEALVINGRIPKPHLPVEREITWQWYQWVTWIPQWDFV 600
 DB 561 KNTHTNGYRLLAHVQIKGKEAIVINGQVPKPHLPVEKDVQWQWTDYWGTYWIPWDFI 620
 QY 601 STPLVRLAPNLVDPFGAETFTYDGCNRSQKQKAGYVTDGKDKVKKLEQTTNQQA 660
 DB 621 STPLVRLVFNLVKDPTEGEBETYYVDGSCQKQKAGYVTDGKDKVKKLEQTTNQQA 680
 QY 661 ELEAFAMALTDSPKVNIIIVDSQYVMGIGISASQPTSESKIVNQIIIEEMIKKEAIIYAVVP 720
 DB 681 ELEAFMALTDSPKANIIVDSQYVMGIIITGCPTESESLVNOIIIEEMIKKEIYVAVVP 740
 QY 721 AHKGIGGNORVDHLVSGIQVQLFLEKIPPAQBEHEKYHSNKKELSHKFGIPNLVARQIV 780
 DB 741 AHKGIGGNDEIDLHVSGIQVQLFLEKIPPAQBEHEKYHSNKKELVFKGLPRLVAKQIV 800
 QY 781 NSCAQCOQKEATHGQVNAELGTWQMDCTHLEKIIIVAVHVASGFIEAEVQESGROT 840
 DB 801 DTCDCHQKEAITHGQVNSDLGTWQMDCTHLEKIIIVAVHVASGFIEAEVQETGROT 860
 QY 841 ALFLLKLASRPITLHTONGANFTSOEVKMWAWTIGQSTGVYVNPQSGVVEAMNHH 900
 DB 861 ALFLLKLASRPITLHTONGANFTSOEVKMWAWTIGQSTGVYVNPQSGVVEAMNHH 920
 QY 901 LKNQISRIEQAANTIEIVLMAITHCMFNRRGGIGDMTPSERLINMITTEQETQFLQAKN 960
 DB 921 LKNQIDRIEQAANSVETIVLMAVHCNRRGGIGDMTPAERLINMITTEQETQFOQSKN 980
 QY 961 SKLKDFRVYREGDQLWGPGLLWKGAGVLVVKVGTDIKIIPRKAIIIRDYGGQEM 1020
 DB 981 SKPFVRYVYREGDQLWGPGLLWKGAGVLVVKVGTDIKIVVPRKAKIIRDYGGQEM 1040
 QY 1021 DSGSHLEGAREDEMA 1036
 DB 1041 DSSSHMEDIGAREVA 1056
 RESULT 13
 AAB12993
 ID AAB12993 standard; Protein; 1055 AA.
 XX AAB12993;
 AC AAB12993;
 XX AAB12993;
 DT 30-NOV-2000 (first entry)
 XX HIV-2 reverse transcriptase protein sequence SEQ ID #39.
 DE HIV-2 reverse transcriptase protein sequence SEQ ID #39.
 XX Reverse transcriptase; RT; polymerase chain reaction; PCR; retrovirus;
 KW stability; solubility.
 XX Human immunodeficiency virus type-2.
 OS Human immunodeficiency virus type-2.
 XX W0200042199-A1.
 XX 20-JUL-2000.
 PD 20-JUL-2000.
 XX 14-JAN-2000; 2000WO-US00896.
 PF 14-JAN-2000; 2000WO-US00896.
 XX 15-JAN-1999; 99US-0116099.
 PR 15-JAN-1999; 99US-0116099.
 XX

(MOLE-) MOLECULAR BIOLOGY RESOURCES.
 Swaminathan N;
 WPI; 2000-482830/42.
 Novel genes encoding reverse transcriptase polypeptides modified by altering or adding the integrase domains by truncation internally and/or at the C-termini, useful in cDNA synthesis and amplification procedures -
 Claim 1; Page 142-146; 189pp; English.
 This invention relates to isolated polynucleotide sequences encoding a polypeptide with RNA dependent DNA polymerase (reverse transcriptase RT) activity. RTs are found in a variety of retroviruses and their defining activity is the ability to synthesise a cDNA strand using an RNA template. The invention includes sequences AAA87808-A87822 and cDNA AAA87840-A87842 which represent nucleotide sequences encoding RT polypeptides. Included in the invention are reverse transcriptase protein sequences AAB12989-B12995. Sequences AAA87829-A87839 and AAA87843-A87897 represent oligonucleotides used in the identification and synthesis of the RT nucleotide sequences of the invention. Many of the RT nucleotide sequences encode modified RT proteins, which exhibit improved stability and/or improved solubility, relative to naturally occurring reverse transcriptases. The modified RT nucleotide and protein sequence are used in improved polymerase chain reaction (PCR) methods. The nucleotide sequences can be used in sequencing methods.
 The present sequence represents a reverse transcriptase protein of the invention.
 Sequence 1055 AA;

Query Match 82.9%; Score 4519.5; DB 21; Length 1055;
 Best Local Similarity 79.7%; Pred. No. 0;
 Matches 839; Conservative 93; Mismatches 86; Indels 35; Gaps 3;
 QY 1 TGRFFRTGPGKEAPQLPR-----GPS-----SAGADTNSTPSSSGSTGEI 43
 DB 18 TGGFFRVRLGKEASQFPFGTGGSAICAPDEPSIRHDTSGDCTICPCRSSRGDAKEL 77
 QY 44 YAAREKTERAERETQGGDRGLTAPRAGDITQATNGLAAPQSLKRPVVTAYIBGO 103
 DB 78 HATREAEQRETEQLQGGD-----RGFAAPQSLMRPVRVAKATIG 119
 QY 104 PVEVLDTGADDSIVAGIELGNYSKIVGGIGGFINTKEYKNVEIYVNLKVRATIMTG 163
 DB 120 SVEVLDTGADDSIVAGIELGNYSKIVGGIGGFINTKEYKNVEIYVNLKVRATIMTG 179
 QY 164 DTPINIFGRNIIALTALGMSLNLPAKVEPIKMLPKGDKGPKLRWPLTKERKEALKEICE 223
 DB 180 DTPINIFGRNIIALTALGMSLNLPAKVEPIKMLPKGDKGPKLRWPLTKERKEALKEICE 239
 QY 224 KMEKEGQLEAEAPPTNPYNTPTFAIKKKDKNKRMLIDFRELNVQDTFETQLGPHPAG 283
 DB 240 KMEKEGQLEAEAPPTNPYNTPTFAIKKKDKNKRMLIDFRELNVQDTFETQLGPHPAG 299
 QY 284 LAKKRITVLDVGDAYFSIPLHEDFRPYTAFTLPSVNNAPGKRIYIYVLPQGWKGSPI 343
 DB 300 LAKKRITVLDVGDAYFSIPLHEDFRPYTAFTLPSVNNAPGKRIYIYVLPQGWKGSPI 359
 QY 344 FOHTMRQVLEPFRKANKDVIIOYMDILLASDRDLEHDRVVLQKLELNGLFSTPDE 403
 DB 360 FOYTMKVLDPFRKANNDVTIIQYMDILLASDRDLEHDRVVLQKLELNGLFSTPDE 419
 QY 404 FKQKDPYHMMGYELWPVKKLOKIQLPQKEITVNDIOKLGVLNWAAQYFGIKTKHL 463
 DB 420 FKQKDPYHMMGYELWPVKKLOKIQLPQKEITVNDIOKLGVLNWAAQYFGIKTKHL 479
 QY 464 CRLIRGKMTLTERVQWTELAELEENRIILSFOEGHYQEEKELEATVQKDOENQWTY 523
 DB 480 CRLIRGKMTLTERVQWTELAELEENRIILSFOEGHYQEEKELEATVQKDOENQWTY 539

QY 524 KIHQEEKILKVGKYAKVKNTHNGIRLLAQQVQKIGKEALYIWIWGRIPKPHLPVEREIQEW 583
 Db 540 KIHQGGKILKVGKYAKVKNTHNGIRLLAQQVQKIGKEALYIWIWGRIPKPHLPVEREIQEW 599
 QY 584 WWDNYWQVWTPIDWDFVSTPLVRLAFNLVGDPIPGAETFFYDGSNCRQSGEGKAGYVTD 643
 Db 600 WMTDYWQVWTPIDWDFVSTPLVRLAFNLVGDPIPGAETFFYDGSNCRQSGEGKAGYVTD 659
 QY 644 RGKDKVKKLEQTTNQAALEAFAMALTDGSGPKVNIIVDSQVYMGISASQPTSESKIVNQ 703
 Db 660 RGKDKVKKLEQTTNQAALEAFAMALTDGSGPKVNIIVDSQVYMGISASQPTSESKIVNQ 719
 QY 704 IIEEMIKKEAIYVAVPAHKIGGQVNDHLVSGIRQVLFLEKIEPAQEEHEKYHSNVK 763
 Db 720 IIEEMIKKEAIYVAVPAHKIGGQVNDHLVSGIRQVLFLEKIEPAQEEHEKYHSNVK 779
 QY 764 ELSHKEGIPNLVARIQVNSCAQCKGEAIGHQVNAELGTWQMDCTHLEGKIIIVAVHVA 823
 Db 780 ELVHKFGIPQLVARIQVNSCAQCKGEAIGHQVNSSELGTWQMDCTHLEGKIIIVAVHVA 839
 QY 824 SGFIEAEVPOESGRQTALFLILKLASRPITHLHTDNGANFTSQEVKMWVWVWIGIEQSGF 883
 Db 840 SGFIEAEVPOESGRQTALFLILKLASRPITHLHTDNGANFTSQEVKMWVWVWIGIEQSGF 899
 QY 884 VYPNPQSGVVEAMNHHKNOISRIEQANTETIVLMAIHCNMFRRGGIGDMPSPRL 943
 Db 900 VYPNPQSGVVEAMNHHKNOISRIEQANTETIVLMAIHCNMFRRGGIGDMPSPRL 959
 QY 944 INMITTEQEIQFLOAKNSKLDVRYFREGDOLWKGPGELLWKGEGAVLVKVGTDIKII 1003
 Db 960 INMITTEQEIQFLOAKNSKLDVRYFREGDOLWKGPGELLWKGEGAVLVKVGTDIKII 1019
 QY 1004 PRRAKIIIRYGGGROEMDSGSHLEGAREDEGMA 1036
 Db 1020 PRRAKIIIRYGGGKELDCSADVEDTMOAREVA 1052
 RESULT 14
 ID AAB01769
 XX AAB01769 standard; Protein; 1018 AA.
 AC AAB01769;
 XX
 DT 03-JAN-2001 (first entry)
 XX
 DE Simian immunodeficiency virus SIVrcm Pol protein #1.
 XX
 KW Red capped monkey; simian immunodeficiency virus; SIV; SIVrcm;
 KW vaccination; Pol protein.
 XX
 OS Simian immunodeficiency virus.
 FH
 FT Key Location/Qualifiers
 FT Misc-difference 569
 FT /note= "encoded by GACTAA"
 XX
 XX WO200034529-A1.
 XX
 XX 15-JUN-2000.
 XX
 PF 03-DEC-1999; 99WO-US28638.
 XX
 PR 07-DEC-1998; 98US-0206551.
 XX
 PA (UABR-) UAB RES FOUND.
 XX
 PI Hahn BH, Shaw GM, Gao F, Marx PA, Smith SM, Georges-Courbot MC;
 PI Lu CY;
 XX
 DR WPI: 2000-423454/36.
 DR N-PSDB: AAA51008, AAB01775, AAB01776.
 XX
 PT New polypeptide, useful for producing antibodies and for diagnosis of

PT simian viral infection in humans comprises complete genome of new
 PT simian immunodeficiency virus isolate -
 XX
 PS Disclosure; Page 152-156; 173pp; English.
 XX
 CC The present sequence is the Pol protein from the simian immunodeficiency
 CC virus found in the red capped mangabey and designated SIVrcm. This virus
 CC is related to the HIV viruses that cause AIDS in humans. Knowing the
 CC sequence of the Pol protein and its gene is useful as it enables
 CC screening for the virus to take place in humans and animals (there is a
 CC possibility that this virus undergoes cross-species transmission), allows
 CC vaccines to be produced and aids research into the origin of these
 CC viruses. This latter is important if an AIDS vaccine is to be found.
 XX
 SQ Sequence 1018 AA;
 Query Match 61.0%; Score 3324; DB 21; Length 1018;
 Best Local Similarity 59.5%; Pred. No. 4; le-258;
 Matches 627; Conservative 156; Mismatches 209; Indels 62; Gaps 8;
 QY 4 FRTGTLGK-EAPOL-----PRGPSSAGADTNSTPSSGSGST 40
 Db 1 FRETPLSGGEAEELSFDFLSCSRDGEQLPCRRDAKELSEERGKETE----- 53
 QY 41 GEIYAAREKTERAERETIQQSDRGLTAPRAGDITQAGATNRGLAAPQFSMLKRPVYAI 100
 Db 54 ----AGREER-----GSISGSLNLPQFALWKRPTTIAQI 85
 QY 101 EGQPEVLLDTCADDISIVAGIELGNYSKPIVGGIGGINTKEYKNVIEVLNKKVATY 160
 Db 86 EGQKEVLLDTCADDITVIEGIELGNDWTPKIIGIGGYINVKYKNCIEIAGKRTAHV 145
 QY 161 MTGDTPIINFGRNILTALGSLNLPVAKVEPTIKMLKPKGDKGPKLROMPLTKETALKE 220
 Db 146 LVGTPYNIIGRNLLKLGATLNFPSIQETIKVELKSGQDGPVRYQVNLPGWKG 205
 QY 221 ICEKMEKEGQLEAPPTNPTPTPAIKKKDKNKRMLIDFRELNKVQDTEIOLGIPH 280
 Db 206 ICNAMEKEGKISKIGPENPTPIFCIKKOSTKWKLVDFRELNKRQDFFEVOLGIPH 265
 QY 281 PAGLAKKRITVLDVGDAYFSIPLHEDFRPTAFTLPSVNAEPKRYIYKVLPGWKG 340
 Db 266 PGLKQCEIRITVLDIGDAYFSCLLYEPFRKYTAFTIPAVNNGQGPVRYQVNLPGWKG 325
 QY 341 PAIFQHTMROVLEPRKANKDVIIQYMDLILIASDRTDLEHDRVVLQKELLNGLGFT 400
 Db 326 PAIFQASANKILQPPREENPDVYIYQYMDLILIASDRTDLEHDRVVLQKELLNGLGFT 385
 QY 401 PDERFKQDPPYHMGYELWPTKWKIQIQLPQKEITWTVNDIQKLVGLNWAALYLPKGT 460
 Db 386 PDKKFDQKPPYLLWGYELHLPKSWTVQEIKLPEKEEWTVDIQKLVGLNWAALYLPKGT 445
 QY 461 KHLCLIRGKMTLPEEVOWTELAELAEENRIILSQOEGHYQYEEKELEATVQKQDNQ 520
 Db 446 KELCLIRGKALDEKVENTEAEIYEENKMLKELHGVYDYDEKPLVANIQLGEGQ 505
 QY 521 WYKITHQEE-KILKVGKYAKVKNTHNGIRLLAQQVQKIGKEALYIWIWGRIPKPHLPVERE 579
 Db 506 WSYQIEQSGPLKTKGAKOKTANTNEIRMLAGLVQKIAKEAIVWGLRFTFRPLPIERE 565
 QY 580 IWEQWMDNYWQVWTPIDWDFVSTPLVRLAFNLVGDPIPGAETFFYDGSNCRQSGEGKAG 639
 Db 566 WVD-WRSQYWQVWTPIDWDFVSTPLVRLAFNLVGDPIPGAETFFYDGSNCRQSGEGKAG 624
 QY 640 YVTORGDKVKKLEQTTNQAALEAFAMALTDGSGPKVNIIVDSQVYMGISASQPTSESK 699
 Db 625 YVTNRGKVKLEQTTNQAALEAFAMALTDGSGPKVNIIVDSQVYMGISASQPTSESK 684
 QY 700 IVNQIIEEMIKKEAIYVAVPAHKIGGQVNDHLVSGIRQVLFLEKIEPAQEEHEKYH 759
 Db 685 LVTEIINQIMIGKEAVYLSWPAHKIGGQVNDHLVSGIRQVLFLEKIEPAQEEHEKYH 744
 QY 760 SNVKSLSHKGIPNLVARIQVNSCAQCKGEAIGHQVNAELGTWQMDCTHLEGKIIIVA 819

XX Sequence encoded by LAV MA L POL gene.
DE HIV; HTLV III; AIDS; diagnosis; vaccine; probe; hybridisation.
KW Lymphadenopathy associated virus MA L.
XX W08707906-A.
XX 30-DEC-1987.
XX 22-JUN-1987; 87WO-EP00326.
XX 23-JUN-1986; 86EP-0401380.
XX (INSP) INST PASTEUR.
PI Allison M, Sonigo P, Wainhobson S, Montagnier L, Allison M;
PI Sonigo P, Wainhobson S, Montagnier L;
XX WPI; 1988-014396/02.
XX N-PSDB; AAN80437.
XX New variants of lymphadenopathy associated virus (LAV) -
PI used for prodn. of DNA, antigens and antibodies used in
PI diagnosis of AIDS and pre-AIDS
XX Claim 8; Fig 8A-8I; 72pp; English.
XX LAV EL I (AAN80436) and LAV MA L (AAN80437) were isolated from the
XX peripheral blood lymphocytes of patients. Different AIDS virus isolates
XX concerned are designated by 3 letters of the patients name. Stable probes
XX including the DNA sequences can be used for detection of the new LAV
XX viruses or related viruses or DNA proviruses in eg. biological samples.
XX The proteins or peptides can be used for detection of antibodies induced
XX in vivo and present in biological fluids. The DNA can also be used for
XX the expression of LAV viral antigens for the prodn. of a vaccine against
XX LAV. The polypeptides can also be used for the prodn. of antibodies for
XX the detection of proteins related to the LAV viruses, partic. for
XX diagnosis of AIDS or pre-AIDS.
XX (Updated on 25-MAR-2003 to correct PI field.)
XX Sequence 1002 AA;

Query Match 57.9%; Score 3157; DB 9; Length 1002;
Best Local Similarity 57.3%; Pred. No. 1.1e-244;
Matches 576; Conservative 177; Mismatches 236; Indels 16; Gaps 5;

Qy 32 PSGSSSGSTGIYAAREKTERAERETIOGSDRGLTAPRAGDGTIOGATNRLAA---POF 88
Db 9 PQGKAREFPSSQTRANSPTSREL- -VMGGDKTLSE-----TGAERQGIVSFPQI 58
Qy 89 SLWKRPPVVTAYIEGQPEVLLDTGADDSIVAGIELGNNSPKYIVGGIGGFINTKEYKNVE 148
Db 59 TLMQRPPVVTYVGGQLEALDLDGADDTVLEELNPKGKWKPKMGIGGIFKVRQYQOIL 118
Qy 149 IEVLNKRVPATIMTGDPIINIFGRNLTALGMSNLVPKVEPIKMLKPGKDGPKLRQW 208
Db 119 IEICGKKAIGTILVGPVNIIGRNMLTQIGCTLNFPISPIETVPVKLPGMDGPRVKQW 178
Qy 209 PLTKEKEALKEICEKEKEGLEEAPPNTDPTPTFAKKKDKNKRMLIDFRELNNVT 268
Db 179 PLTEEKIKALTEICKMEKEGKILKIGPENPNTVPFAIKKDKSTKWRKLVNRELNKT 238
Qy 269 QDFTIOLGPHPLAGLAKKRITVLDYGDVAFSLPHEDFRPYTAFTLPVNNAPGKRY 328
Db 239 QDFVEVLGPHPLAGLAKKSVTLVDYGDVAFSVPLDEDFPKYTAFTIPSINNETPGIRY 298
Qy 329 IYKVLPGWKGSPAIFQHTMRQVLEPFRKANKDVIIIIQYMDIILIASDRTDLEHDRVQLQ 388
Db 299 QYNVLPQWKGSPAIFQSSMTKILEPFRKNTKPEIYIYQYMDLVGSDLEIGHRTKLEE 358
Qy 389 LKELLNGIGFSTPDEKFKDPPYHWMGYELWPTKWKLOKIQLPQKEIWTVNDIQKLVGL 448

Db 359 LREHLKMGFTTDPDKKHOKEPFLWMGYELHPDKWTQPIQLPKESWTVDIQKLVGL 418
Qy 449 NWAOLYPGIKTKHLCLRIKGMKMTLTFEEOVQTELAELAEENRIILSQEGHYQBEKE 508
Db 419 NWSOIYPGIKVKQLCKLRGAKALTDIVPTAEAELELPENRELKEFVGVVYDPSKD 478
Qy 509 LEATVQKQDNQWYKIHQEE- KILKVGKYAKVKNTHNGIRLLAQVYQKIGKALVIWG 567
Db 479 LIAETQKOGQWYQIYEQYKNLTKGYARIKSAHTNDVKQLTEAVQKIAQESIVIWG 538
Qy 568 RIPKPHLPVEREINWQWWDNWTWIPDWDFVSTPPPLVRLAFNLVGDPIPGAETFYDQ 627
Db 539 KTKPRLPIQKETWEAWWTEYQATWIPWEFVNTPLVKLWYQLETEPIVGAETFYDQ 598
Qy 628 SCNRQSGKAGYVTDGKDKVKLEOTTNQOALEAFAMALTDSGPKVNIIVDSOYVMG 687
Db 599 AANRETKKAGYVTDGRQKRVSLTETTNKTELQAIHLALQDSGSEVNIIVDSQYALG 658
Qy 688 ISASOPTESKIVNOIIEEMIKKEAIYVAVWPAHKIGGIGNOEVDHLYSOGIROVLFLEK 747
Db 659 LIQAPDKSESEIVNQIIEQLIQDKYLSWVPAHKIGGIGNOEVDKLYSSGIRKVLFDG 718
Qy 748 IEPAQEEHEKYHSNVKLSHKGIPNLVAROIVNSCAQCKQKGAIBHGVNAELGTWMD 807
Db 719 IDKAQEEHEKYHSNWRAMASDFNLPPIVAKETIVASCDKQLKGEAMHGQVDCSPQIMOLD 778
Qy 808 CTHLEGGKTIIVAVHVASGFIKAEVTPQESGRTALFLKLLASRPITHLHTDNGANFTSQ 867
Db 779 CTHLEGGKTIIVAVHVASGYIEAEVTPAETGTETFAIFLKLGRWPKVYVHDNGSNFTSA 838
Qy 868 EVKRVAMWIGIEQSGFVPYNPQSQGVVAMNHLKNQISRIREQANTTETIVLMAIHCNM 927
Db 839 AVKAACWANIKQEFIPYNPQSQGVVSMKELKLIIGOVREQAEHLKTAQVQMAVFIHN 898
Qy 928 FKRGGIGDMTPSERLINMITTEBIEQIQLAKNSKLKDFRYVYFREGDQLWKGEGELLWK 987
Db 899 FKRGGIGGYGAGERIIDMIATDITQKELQKIQKIONFRVYRDNRDPIWKGPAKLWK 958
Qy 988 GEGAVLVKVGTDIKIIPRKAIIIRDYGGROEMDSGSHLEGARE 1032
Db 959 GEGAVVIQDNSDIKVPPRRKAKIIRDYG--KOMAGDCCVAGGQDE 1001

RESULT 17
AAW72993
ID AAW72993 standard; Protein; 1002 AA.
XX
AC AAW72993;
DT 25-MAR-2003 (updated)
DT 15-FEB-1999 (first entry)
XX
XX HIV isolate LAV.MAL pol protein.
DE LAV.MAL; HIV; human immunodeficiency virus; diagnosis; vaccine;
KW AIDS; pol.
XX Lymphadenopathy-associated virus isolate MAL.
OS
XX US5824482-A.
PN
XX 20-OCT-1998.
PD
XX 06-JUN-1995; 95US-0471474.
PE
XX 23-JUN-1986; 86FR-0040138.
PR 13-APR-1987; 87US-0038330.
PR 19-FEB-1991; 91US-0656797.
PR 10-DEC-1992; 92US-0988530.
PR 18-NOV-1993; 93US-0154397.
XX
XX (INSP) INST PASTEUR.

Seq	Sequence	1014 AA;
	Query Match	57.8%; Score 3150.5; DB 19; Length 1014;
	Best Local Similarity	56.0%; Pred. No. 3.8e-244;
	Matches	579; Conservative 179; Mismatches 253; Indels 23; Gaps 6;
Qy	4 FFRGTGPGKAPQLPRGPSSAGADNTSPSSSGSTGEIYAAREKTERAERETIQ-GSD 62	
Db	1 FFR-----EELVSLQRETRKLPDPNNK--EHAHSPATRELWVSGGEHTGEGDAGEPGED 53	
Qy	63 RGLTAPRAGGDTIQGATNRGLAAPQFSLWKPVPVYATIEGQPVVEVLDTGADDSIVAGIE 122	
Db	54 RELSVPT-----FNFQITLWQRPVITVKIGKEVREALLDTGADDTVIELQ 100	
Qy	123 LGNNYSPKIVGIGGIFNTKEYKNVEIEVLNKKVRATIMTGDTPINIFGRNLTALGMSL 182	
Db	101 LEGRWPKMIGGIGGIFKVRQYDNTDIQKQKAVGTIVLGPVPTVNIIGRNLLTQIGCTL 160	
Qy	183 NLPVAKVEPIKMLKPGDKGPKLRQWPLTKKIEALKEICEKMEKEGQLEAPPTNPNYNT 242	
Db	161 NFPISPIETVPVKLPKPGMDGPKVQWPLTTEKIEALREICTEMEKEGKISRIGPENPNYNT 220	
Qy	243 PTEAIKKKKDKNKKWMLIDFRELKVKWTDQFTETIQLGPHIPAGLAKKRRITVLDVGDAYESI 302	
Db	221 PIFAIAKKKDSFKWKKLVDFRELKRTQDFWEVQLGPHIPAGLKQKSVTVLVDVGDAYEFC 280	
Qy	303 PLHEDFRPYTAFTLPSVNNAPGKRYIYKVLPGQWKGSPAIFQHTMROVLEPFRKANKDV 362	
Db	281 PLDKDFRYTAFTPSINNETPGIRYOYNVLPGQWKGSPAIFQSTMTKILFFREKHPFI 340	
Qy	363 IIIQYMDLILASBRTDLHRVVLQKELNGLGTFSPDBKFKQDPPYHWMGLMPTK 422	
Db	341 IIYQYMDLYVGSLSLELAHQREAVEDRLHLKMGFTTPDKKHQKEPPFLWMGYELHDPK 400	
Qy	423 WKLOKIQLOPEIKTWNVDIOKLVGLNWAQOLYPGIKTKHLCLRLGKNTLITEVQWTEL 482	
Db	401 WTVQIKLPEKDWMTVNDIOKLVGLNWAQOLYPGIRVKQJCKLIRGARALTVEYVNFTEE 460	
Qy	483 AEAELENRIILSQEGBHYHQEKELEATVQKQENQWTKYKHQ--EKILKVGKYAKVK 541	
Db	461 AELEAENREILKEPLHGVYDDPGKELVAELQKOGQWTVYIYQELHKNLKTGYAKMR 520	
Qy	542 NHTNGIRLLAQVYQKIGKELVNIWGRIPKPLHVERIWFQWMDNYSWQVWIPWDFVS 601	
Db	521 SAHTNDIKOLVEVYRKVATESIVIGWTKPRFLPVQKEVWDAWNTDHWQATWIPWEFVN 580	
Qy	602 TPLVRLAFNLVGPDPGAETFTYDGSNRSKESKAGYVTRDCKVKKLEQTNQOAE 661	
Db	581 TPLVRLKWLQLETFISGAETFTYDGAANRHLKLGAGFVTRDGRQKVVSTADTNNQAE 640	
Qy	662 LEAFAMALTDGPKVNIIVDSQYVNGISASOPTSESKIVNQIIEEMIKKEAIYVAVWPA 721	
Db	641 LQAILMALQESGRDNIIVTDSQYAMGIHSPDKSESELVSQIIEELIKKERVYLSWVPA 700	
Qy	722 HKGIGGNOVDHLVSGIRQVLFLEKTEPAQEEHEKHYSNVKLSHFGIPNLVARQIVN 781	
Db	701 HKGIGGNEQVDKLVSSGIRKILFLDIEKAQEDHRYHSNWKAMASDFNLPIVAKELVA 760	
Qy	782 SCAQCQKGAIRHGSVNAELGTWQMDCTHLEGLIIVAVHVASGFIAEVLIPQESGRQTA 841	
Db	761 SCDCQKLGEMHGSQVNCSPGVQWLDCTHLEGLIIVAVHVASGFIAEVLIPQESGRQTA 820	
Qy	842 LFLKLASRPITHLHNDGANFTSOEVKVMVWIGIEQSGFVPYNPQSQGVVAMNHL 901	
Db	821 YFILKLAGRVPVKVIHTDNGSNFTSATVKAACWANIKEQGIYPNPQSQGAVESMKNEL 880	
Qy	902 KNOISRTREANTTETIVLMAIHCNPKRRGGIDMTSPERKLINMITTEQIOFLQAKNS 961	
Db	881 KKIIGQIRDOAEHLKTAQVMAVFTHNFKRKGIGGYTAGERIIDIATIDQITNLQTOIL 940	
Qy	962 KLKDFRYFERGRDOLWKGPCELLWKGEGAVLVKVGTDIKIIPRKKAKIIRDYGRQEMD 1021	
Db	941 KVQNFVRYRSDRPIWKGPAKLLWKGEGAVVIQDNGDIKVVPRRKAKIIRDY--GKQWAG 999	

Qy	1022 SGSHLEGAREDGEM	1035
Db	1000 DGCVASGQDENQEM	1013

RESULT 19
AAB69287
ID AAB69287 standard; Protein; 1005 AA.
XX AAB69287;
AC AAB69287;
XX 20-APR-2001 (first entry)
XX HIV-1 non-subtype B clone 962M751-3 pol protein.
DE HIV-1; human immunodeficiency virus; non-subtype B; gag; pol; env;
KW vpu; vif; vpr; tat; rev; nef; vaccine.
XX Human immunodeficiency virus type 1.
OS
XX WO2000026416-A1.
PN
XX 11-MAY-2000.
XX 25-OCT-1999; 99WO-US24837.
XX 02-NOV-1998; 98US-0184418.
XX (UABR-) UAB RES FOUND.
PA
XX Hahn BH, Shaw GM, Gao F;
PI
XX WPI; 2000-365651/31.
DR
XX Novel genomic nucleic acids of non-subtype B human immunodeficiency virus type 1 useful for detecting and treating AIDS comprises a specific nucleotide sequence
PT
XX Claim 41; Fig 15; 131pp; English.
PS
XX The present invention provides the protein and coding sequences for a number of human immunodeficiency virus (HIV) type 1 non-subtype B isolates. The sequences shown include the near full-length coding sequences from each isolate, and the env, pol, vif, vpr, vpu, gag, tat, rev and nef proteins. These can be used to detect the presence of HIV-1 in a sample and to produce antibodies against non-subtype B HIV-1 virus. These antibodies can be used in vaccines to prevent and treat HIV infection.

Seq Sequence 1005 AA;

Query Match 57.5%; Score 3134.5; DB 21; Length 1005;
Best Local Similarity 55.9%; Pred. No. 7.3e-243;
Matches 577; Conservative 184; Mismatches 230; Indels 41; Gaps 9;

Qy	4 FFRGTG---PLGKEAPOLPRGPSSAGADNTSPSSSGSTGEIYAAREKTERAERETIQ 60	
Db	1 FFRLENLAPPEG-EGALPSEQTRA-----NSPTSSNSPTREL-----QVRG 41	
Qy	61 SDRGLTAPRAGGDTIQGATNRGLAAPQFSLWKPVPVYATIEGQPVVEVLDTGADDSIVAG 120	
Db	42 DN---PCPEAGAER-QGTFLN---CPQITLMORPLVSIKVGQIKREALDITGADDTVLEE 93	
Qy	121 IELGNYSPPKIVGGIGGIFNTKEYKNVEIEVLNKKVRATIMTGDTPINIFGRNLTALGM 180	
Db	94 INLPCKWPKMIGGIGGIFKVRQYDQILIEICGKKAIGTVLVGPVPTVNIIGRNMLTQJGC 153	
Qy	181 SLNLPVAKVEPIKMLKPGDKGPKLRQWPLTKKIEALKEICEKMEKEGQLEEAAPTNPY 240	
Db	154 TLNFPISPIETVPVKLPKPGMDGPKVQWPLTTEKIEALKEICEKMEKEGKTKIGPENPY 213	
Qy	241 NPTPTAIKKKKDKNKKWMLIDFRELKVKWTDQFTETIQLGPHIPAGLAKKRRITVLDVGDAYF 300	

```
Db 214 NTPVFAIKKSDTKWRKLVDFRELNKRQDTEVEVQLGIPHPAGLKKKSVTVLVDGDAYF 273
Qy 301 SIFLHEDFRPYTAFTLPSVNNAPGKRYIKVLPQGWKSGSPAIFQHTMRQVLEPRKANK 360
Db 274 SVPLDEGFRKYTAFTLPSINNETPGIRYQYNVLPQGWKSGSPSIFQSSMIKILEPFTQNP 333
Qy 361 DVIIQYMDLIIASDRTDLEHDRVVLQKELNGLGFGSTDPKFKDPPYHWMGYELWP 420
Db 334 EIVIIQYMDLIVGSDLEIGHRAKTEELREHLRWGFTTPDKKHQKEPFFLWMGYELWP 393
Qy 421 TKWKIQIOLPOKEIWTNDIOKLVGLVNWAAQLYPGIKTKHLRLRIGKMTLIEEVQWT 480
Db 394 DKWTVPQIKLPEKESVTNDIOKLVGLKNWSQIYAGIKVRQCKLLRGAKALTDIVPLT 453
Qy 481 ELAAEAELEENRIILSQEKGHYQBEKELEATVOKDOENQWYKIHQEE-KILKVGKYAK 539
Db 454 EAAELEABSEIREILKEPVGHYVYDPSKDLIAEQKGHDWYQVYQVEPFFKLTGKYAK 513
Qy 540 VKNHTNGIRLLAAQVQKIGKCALVIWGRIPFHLPPVEREIVEQWMDNTWQVTWIPDWF 599
Db 514 MRTAHTNDVKQLTEAVOKTAMESIVLWGIKIPERLDIQKETWETWTDIWAQATWPEWF 573
Qy 600 VSTPPLVRLAFNLVGDPIPAETFTYDSCNRQSKGKAGYVTDGRKDKVKVKEQTNQO 659
Db 574 VNTPPLVLKWLQLEKEPIAGAETYYVDGAANRETKIGKAGYVTDGRQKIVTLTETNOK 633
Qy 660 ALEAFAMALTDGPKVNIIVDSQYVMGISASQTESKTYNQIIEEMIKKEAYVAVW 719
Db 634 TELQATLALQDGSSEVNTVTSQYALGIIQAQPKSESELVNIQIEOLIKERVYLSWV 693
Qy 720 PAHKGIGGNOEVDHLVSGIROVLFLEKTEPAQEHEKHYSHNVKELSHKFGIPNLVAROI 779
Db 694 PAHKGIGGNEQVDKLVSSGIRKVLFDGIDKAEHEKHYSHNWAMASEFNLPPIVAKEI 753
Qy 780 VNSCAQCOQKGAIRHGQVNAELGTWQMDCTHLEGKIIIVAVHVASGFIEAIVPOESGRQ 839
Db 754 VASCDKQLKGAIRHGQVDCSPGIWQDCTHLEGKVLVAVHVASGYIEAIVTPAETGOE 813
Qy 840 TALFLLKLASRPFILTHLDNGANTSOEYKVVANWIGIEQSGVPYNPQSGOVVEAMNH 899
Db 814 TALFLLKLGRPVKVVHHDNGSNFTSAAVKACWAGIHQEGIPYNPQSGOVVESMKN 873
Qy 900 HLKNOISIREQANTIETIVLMAHCMNFKRRGGIGDMTPSRRLNMIITEOEIOFLQAK 959
Db 874 ELKKIIGQVRDQAEHLKTAQVAVFTHNFKRGGIGGYSAGERIIDIIATDIOTRELQKQ 933
Qy 960 NSKLDKFRVYFREGDQLWKGCELWKGEGAVLVKVGTDIKIIPRKAIIIRDYG---- 1015
Db 934 IIKIQNFYRYRDSRDPWKGPAPKLLWKGEGAVVIQDNSDIKVIIPRKAIIIRDYGKQMA 993
Qy 1016 -----GROEMD 1021
Db 994 GTDSVAGRODED 1005

RESULT 20
AAB69288
ID AAB69288 standard; Protein; 1003 AA.
XX
AC AAB69288;
XX
XX 20-APR-2001 (first entry)
XX
DE HIV-1 non-subtype B clone 94CY017-41 pol protein.
XX
KW HIV-1; human immunodeficiency virus; non-subtype B; gag; pol; env;
KW vpu; vif; vpr; tat; rev; nef; vaccine.
XX
OS Human immunodeficiency virus type 1.
XX
PN
XX
XX WO200026416-A1.
XX
```

```
11-MAY-2000.
XX
XX 25-OCT-1999; 99WO-US24837.
XX
XX 02-NOV-1998; 98US-0184418.
XX
PA (UABR-) UAB RES FOUND.
XX
XX Hahn BH, Shaw GM, Gao F;
DR WPI; 2000-365651/31.
XX
XX Novel genomic nucleic acids of non-subtype B human immunodeficiency
PT virus type 1 useful for detecting and treating AIDS comprises a
PT specific nucleotide sequence
XX
XX Claim 41; Fig 15; 131pp; English.
XX
XX The present invention provides the protein and coding sequences for a
CC number of human immunodeficiency virus (HIV) type 1 non-subtype B
CC isolates. The sequences shown include the near full-length coding
CC sequences from each isolate, and the env, pol, vif, vpr, vpu, gag, tat,
CC rev and nef proteins. These can be used to detect the presence of HIV-1
CC in a sample and to produce antibodies against non-subtype B HIV-1 virus.
CC These antibodies can be used in vaccines to prevent and treat HIV
XX infection.
XX
XX Sequence 1003 AA;
XX
XX Query Match 57.5%; Score 3133; DB 21; Length 1003;
XX Best Local Similarity 58.3%; Pred. No. 9.6e-243;
XX Matches 575; Conservative 164; Mismatches 226; Indels 22; Gaps 6;
Qy 52 RAERET---IQSDRLGTAPRAG-GD--TIQATNGLAAPQPSLMKRPVTVAYIEGQPV 105
Db 22 RANSTPSRELENGGRDNLPEAGTGQGTQISCN----FPQITLWQRPVLTVKIEGQLK 76
Qy 106 EYLLDTGADDSIVAGTELGNNYSPKIVGGIGGFINTKEYKNVEIEVLNKKVRATIMTGD 165
Db 77 EALLDTGADDTVLEELNLPCKWKPMIGGIGGFVKYRQYDQIAIEICGKRAIGTIVLVGPT 136
Qy 166 PINIFGRNLTALGMSLNLPAKVEPIKIMLPGKDGPKLRQWPLTKKEIKALKETCEKM 225
Db 137 PNYIIIGRNMLVQLGCTLNPISPIETVPVKLAPGMDGPKVKQWPLTEERIKALTEICKEM 196
Qy 226 EKEGOLAEAPPNPYNTPTFAIKKDKKRWMLIDPRELNKVTODTEIQLGIPHPAGLA 285
Db 197 EKEGKISKIGPENPNYTPVFAIKKDKSTKRWKLVDFRELNKRQDTEVEVQLGIPHPAGL 256
Qy 286 KKRRTITVLDGDAYFSLPHEDFRPYTAFTLPSVNNAPGKRYIKVLPQGWKSGSPAIFQ 345
Db 257 KKKAVTVLDGDAYFSVPLHEDFRKYTAFTIPTNNETGVRVQYNVLPQGWKSGSPAIFQ 316
Qy 346 HTMROVLEPFRKANKDVIIQYMDLIIASDRTDLEHDRVVLQKELNGLGFGSTDPKFK 405
Db 317 SSMTKILEPFRSKNTELIYQYMDLIVGSDLEISQHRVKIEELRAHLKKGFTYDPDKKH 376
Qy 406 QKDDPPYHWMGYELWPQKWKLOKLOLPKEIWTNDIOKLVGLVNWAAQLYPGIKTKHLGR 465
Db 377 QKEPFFLWMGYELHPDKWTVQPIKUPKEDSVNTNDIOKLVGLKNWASQIYAGIKVQLCK 436
Qy 466 LIIRKMTLIEEVQWTELAEELEENRIILSQEKGHYQBEKELEATVOKDOENQWYKIH 525
Db 437 LLRGAKALTDIVLTKEAELEENREILKTPVHGYYDPSKDLIAEIOKODQOWTYOI 496
Qy 526 HQEE-KILKVGKYAKVKNTHNGIRLLAAQVQKIGKALVIWGRIPKPHLPEREIVEQW 584
Db 497 YOPEPKNLKTGKYAKRRSTHTNDIKQLTEAVQKIMTESIVWGTPTKPKLTIQKETWET 556
Qy 585 WDNVWQVTWIPDWFVSTPPLVRLAFNLVGDPIPAETFTYDSCNRQSKGKAGVTVDR 644
Db 557 WAELYQATWIPWEPFVNTPLVLKWLQLEKEPIAGAETFYVDGAANRETKLKGAGVTVDR 616
```



```
Db 11 GKAREFSEQTRANSPTRRELQVWVGGENSL--EAGADR-QGTVSFNF--PQITLWQRP 65
Qy 95 VVTAYIEGQPVVLLDTGADDSIVAGIELGNYSKPIVGGIGGINTKEYKNVEIEVLNK 154
Db 66 LVTIRIGGOLKEALLDTGADDTVLEEMNLPFGKWKPKMIGGIGGIFKVRQYDQIPVEICGH 125
Qy 155 KVRATIMTGTPTINIFGRNLTALGMSLNLPAKVEPIKIMLPGKDGPKLRQWPLTKRK 214
Db 126 KAIGTVLVGPTPVNIIGRNLLTQIGCTLNFPISPIETVPVKLPQMGMDGPKVQWPLTEBK 185
Qy 215 IEALKEICKMEKEGLEBAPTPNPTPTFAIKKDKNKNRMLIDFRELNVKTDFTFI 274
Db 186 IKALVEICTEMEKEGKISKIGPENPYNTPVFAIKKDKSTKRWKLVDFRELNKRQDQFEV 245
Qy 275 QLGIPHPAGLAKRRITVLVDGDAVESIPLHEDFRPYTAFTLPSVYNAEPGRKIYKVL 334
Db 246 QLGIPHPAGLKKKKSVTLVDGDAVSVPLDKFRKYTAFTIPSTINNETPGRIYQYNL 305
Qy 335 QGWKSGPAIFQHTMRQVLEPFRKANKDVIIIOYMDILLIASDRTDLEHDRVVLQKLELN 394
Db 306 QGWKSGPAIFQSSMKILEPFRKQNPDIYIYQYMDLLVGSDELIGQHRKTEELRQHLL 365
Qy 395 GLGFSTPDKFOKDPYPHWMGYELMPTKWLQKIQLPQKEIWTVDIOKLVGLVNLMAQL 454
Db 366 RWGFTTPDKKHQEPFLMWGYELHPDKWTVQPIMLPEKDSWTVDIOKLVGLKLNWASQI 425
Qy 455 YPGIKTKHLIRLGRKMTLVEVQWTELAELAELENRIILSQEGBHYQEEKELEATVQ 514
Db 426 YAGIKVKQCKLRLGRKTAUTVPIPTEEAELEAENREILKEPVHEVYDPSKDLVAETQ 485
Qy 515 KQENQWTKIHQEE--KILKVGKYAKVKNTHNGIRLLAQVYQVKIGKEALVIGRIPKPH 573
Db 486 KOGQGWTVYIQEFPKNLKTQYARMGAHNDVKQLTEAVQKYSTESIVLWGIKIPRK 545
Qy 574 LPVERIEQWMDNTWQVWIQDVPSTPPVLRLAFNLVGDPIPCAETFFYDGCNQRS 633
Db 546 LPIQKETWAWMNEYQATWIPEWFEVNTPPVLKWLQYQLEKEPIVGAETFFYVDGAANRET 605
Qy 634 KEGKAGYVTRDKGKVKLEOTNQOAELEAFAMALTDSGPKVNIIVDSQYVMGISAPO 693
Db 606 KLGKAGYVTRDGRQKVVSTADTTNQKTELQAHLALQDSGLEVNIVTDSQYALGIIQAP 665
Qy 694 TESKIVNQIIEEMIKKBAIYVAVPAHKGIGGQVNDHLVSOGITROVLFLFLEKIEPAQE 753
Db 666 DKSESELVSOIIEQLIKKEKVLAWYPAHKGIGGQVNDKLVYAGIRKVLFLNGIDKQAE 725
Qy 754 EHEKYSNVKLSHFGIPNLVARQIVNSCAQCKQGEAIGHQVNAELGTWMDCTHLEG 813
Db 726 EHEKYSNWRWASDFNLPPVVAKEIVASCDCQLKGEAMHGQVDCSPGIWLDCTHLEG 785
Qy 814 KIIIVAVHVASGFIIEAVIPEQSGROGTALFLLLKLASRPITHLHTDNGANFTSQEYKMYA 873
Db 786 KIILVAVHVASGYIEAEVIPAETGQETAYFLKLGRWPVKTIHTDNGSNFTSTIVKAAC 845
Qy 874 WWIGIEQSGFYPNQSGVBEAMNHLKNQISIREQANTTETIVLMAIHCMPNKRKG 933
Db 846 WWAGIKQEPGYPNQSGVBEAMNHLKNQISIREQANTTETIVLMAIHCMPNKRKG 905
Qy 934 IGDWTPSELINNIITEQEIQFLOAKNSKLDKFRVYFREGROQLWKGGELLWKGEVAVL 993
Db 906 IGGYAGERIVDIIATDQTKELQKITQIONFRVYRDNDKPLMKGPFAKLWKGEVAV 965
Qy 994 VKVGTDIKIPRKAIRDYGGQEMDSGSHLEGARD 1032
Db 966 IODNSDKVVPARKAKIIRDYQKM---AGDDCVASRQD 1001
```

RESULT 25

AAV77301

ID AAY77301 standard; Protein; 1003 AA.

XX

AC AAY77301;

```
XX 22-MAY-2000 (first entry)
XX HIV-1 (ATCC CRL 8597, single isolate) pol protein.
DE HIV-1; immunoassay; antigen; pol fragment; p31; recombinant; antibody;
XX detection.
KW Human immunodeficiency virus type 1 'ATCC CRL 8597'.
XX
OS Key Location/Qualifiers
FH Misc-difference 1003
FT /note= "Encoded by GAT"
XX
XX US6013432-A.
XX 11-JAN-2000.
XX 17-MAY-1995; 95US-0443434.
XX 08-JUL-1993; 93US-0089407.
PR 24-DEC-1987; 87US-0138894.
PR 17-AUG-1992; 92US-0931154.
PR 31-OCT-1984; 84US-0667501.
PR 30-JAN-1985; 85US-0696534.
PR 06-SEP-1985; 85US-0773447.
XX (CHIR ) CHIRON CORP.
XX
XX Luciw PA, Dina D;
PI WPI; 2000-170256/15.
XX N-PSDB; AA290202.
XX
XX Immunoassay for antibodies against human immune deficiency virus, for
PT diagnosing infection, uses an immunogenic fragment of the pol protein
PT as antigen -
XX
XX Example 1; Fig 5B-E; 99pp; English.
XX
XX The invention relates to the improvement of HIV-1 immunoassays by the
CC use of an HIV-1 antigen comprising an immunogenic fragment of
CC recombinant or synthetic HIV-1 pol, which is encoded by an approximately
CC 9.7 kb sequence between a BstXI restriction site at position 3006 and an
CC NdeI site at position 5131 of the genome (the proviral DNA sequence is
CC given in AA290201). The immunogenic pol fragment is not immunologically
CC cross-reactive with human T cell lymphotropic viruses I or II. The
CC invention also encompasses the use of p31 as an antigen. The recombinant
CC antigens may be produced in Escherichia coli, Saccharomyces cerevisiae or
CC in mammalian cells. Immunoassays using the recombinant HIV proteins may
CC be used to diagnose and stage HIV-1 infections. Sequences
CC AAY77300-Y77302 represent, respectively, the gag, pol and env proteins
CC encoded by the proviral DNA of a single isolate of HIV-1 (ATCC CRL
CC 8597).
```

XX SQ Sequence 1003 AA;

Query Match 57.2%; Score 3119; DB 21; Length 1003;

Best Local Similarity 56.8%; Pred. No. 1.3e-241;

Matches 567; Conservative 179; Mismatches 241; Indels 12; Gaps 6;

Qy 38 GSTGIYAAAREKTERAERETIQ---GSDRGLTAPRAGGDTIQGATNRGLAAPQFSLWKR 94

Db 11 GKAREFSEQTRANSPTRRELQVWVGGENSL--EAGADR-QGTVSFNF--PQITLWQRP 65

Qy 95 VVTAYIEGQPVVLLDTGADDSIVAGIELGNYSKPIVGGIGGINTKEYKNVEIEVLNK 154

Db 66 LVTIRIGGOLKEALLDTGADDTVLEEMNLPFGKWKPKMIGGIGGIFKVRQYDQIPVEICGH 125

Qy 155 KVRATIMTGTPTINIFGRNLTALGMSLNLPAKVEPIKIMLPGKDGPKLRQWPLTKRK 214

Db 126 KAIGTVLVGPTPVNIIGRNLLTQIGCTLNFPISPIETVPVKLPQMGMDGPKVQWPLTEBK 185

```
QY 215 IEALKEICEKMEKEGQLEAEPTNPNTPTFAIKKDKKWKRMILDFRELNKVTQDFTEI 274
Db 186 IKALVEICTEMEKEGKISKIGCPENPTPVFAIKKDKSTKWRKLVDFRELNKRTQDFWE 245
QY 275 QLGIPHPAGLAKKRITVLDVGDYFSLPHEDRPPYTAFTLPVNNAEPCGRYIKVLP 334
Db 246 QLGIPHPAGLAKKRITVLDVGDYFSLPHEDRPPYTAFTLPVNNAEPCGRYIKVLP 305
QY 335 QGWGSPALFQHTMRQVLEPPRKANKDVIITQYMDLIIASDRTDLEHDRVVLQKLKLN 394
Db 306 QGWGSPALFQHTMRQVLEPPRKANKDVIITQYMDLIIASDRTDLEHDRVVLQKLKLN 365
QY 395 GLGSTDEKQKPPHYHMGYELWPTKWKLOKIQLPQKEITWVNDIQKLVGLNWAQAL 454
Db 366 RWGFTTDPKHKKEPPFLMKGVELHPDKWTQVIMLPKDSWTVNDIQKLVGLNWAQAL 425
QY 455 YPGIKTRHLCRLIGKMTLIEVQVOWTELAELLENRIILSQOEBGHYYQKEKELEATVQ 514
Db 426 YAGIKVQOLCKLGRKALTVEIPLTEAELEAENREILKEPHEVYIDPDKSLVARIQ 485
QY 515 KDOENOWTYKTHQEE-KILKVGKAKVKNTHNTINGIRLLAOVVOKIGKEALVIGRIPKFH 573
Db 486 KOGQOWTYQIYQPFNMLKTKGYARMGAHTNDVKQTEAVQKVSTESIVWIKIPKFK 545
QY 574 LPVERIEQWMDNYQVWTIPDWDVSTPLVRLAFNLVGDPIPGAETFTDGCNRSQS 633
Db 546 LPIQETWEAWMEYQWATWIPWFEVNTPLVRLVQLEKEPIVGAETFFVDGAANRET 605
QY 634 KEGAGYVTRDGRKDKVKKLEQTTNQAELEAFAMALTDSPKVINIVDSQVYMGISASQP 693
Db 606 KLGRAGYVTRDGRKDKVKKLEQTTNQAELEAFAMALTDSPKVINIVDSQVYMGISASQP 665
QY 694 TESEKIVNOIIEEMIKKEALYVAVPAHKGIGGNOEVDHLVSGIROVLFLEKTEPAQE 753
Db 666 DKSESELVQIIEQLIKKEKYLAWPAHKGIGGNEQVDKLVSGIRKVLFLNGIDKAEQ 725
QY 754 EHEKHSNVKLSHKFGIPNLVAROIVNSCAOCCOQGEAIIHQVNAELGTWQMDCTHLEG 813
Db 726 EHEKHSNWRMASDFNLPPVVAKEIVASCDCOLKGEAMHQVDCSPGIWQDCTHLEG 785
QY 814 KIIIVAVHVASGFTEAEVPOESGROTALFLKLKASBPITHLHTDNCANETSQEVKMAV 873
Db 786 KIIIVAVHVASGFTEAEVPOESGROTALFLKLKASBPITHLHTDNCANETSQEVKMAV 845
QY 874 WIIIEQSFQYPYNPQSGVVEAMNHLKNOISRIEQANTIEITVLMIAHCMNFKRRGG 933
Db 846 WIIIEQSFQYPYNPQSGVVEAMNHLKNOISRIEQANTIEITVLMIAHCMNFKRRGG 905
QY 934 IGMTPSERLINMTTQEQIQLQAKNSKLDKFRVYFREGDQLWKGPGLLWKGEGAVL 993
Db 906 IGGYSAGERIVDIIATDITQTELQKITKIQNFRVYVRDNKDLPLWKGPGLLWKGEGAV 965
QY 994 VKVCTDIKIPRRKAKIIRDYGGQEMDSGSHLEGARD 1032
Db 966 IQDNSDIKVPVPRKAKIIRDYGGQEMDSGSHLEGARD 1001
```

RESULT 26

AAY70601
ID: AAY70601 standard; Protein; 1003 AA.

XX AAY70601;

XX 18-JUL-2000 (first entry)

DE Corrected version of wild type HIV pol protein.

XX HIV; gag; packaging cell line; lentivirus; retroviral vector particle;
KW pol; gene therapy; gene replacement; vaccine; biochemical reagent.

XX Human immunodeficiency virus.

XX WO200015819-A1.

```
XX 23-MAR-2000.
PD 10-SEP-1999; 99WO-US20675.
XX 11-SEP-1998; 98US-0100022.
PR 12-SEP-1998; 98US-0100063.
XX (CHIL-) CHILDRENS MEDICAL CENT.
PA Gray JT, Mulligan RC;
PI WPI; 2000-271455/23.
XX N-PSDB; AAZ52053.
PT New packaging cell line for producing a viral accessory protein
PT independent HIV derived retroviral vector particles, useful in gene
PT therapy or gene replacement -
XX Disclosure; Fig 9; 62pp; English.
XX The patent discloses new packaging cell line for producing a viral
CC accessory protein independent lentivirus, preferably
CC human immunodeficiency virus (HIV), derived retroviral vector particles.
CC The packaging cell line comprises a mammalian cell, a retroviral DNA
CC comprising a coding sequence for a lentivirus, preferably HIV, gagpol,
CC where the coding sequence has been mutagenised to improve expression of
CC the viral gagpol proteins, a second retroviral nucleotide sequence
CC comprising the coding sequence for a heterologous envelope protein and a
CC third retroviral nucleotide sequence comprising a DNA sequence of
CC interest and lentivirus, preferably HIV, cis-acting sequences required
CC for packaging, reverse transcription and integration.
CC The packaging cell lines and viral particles can be used for gene
CC therapy or gene replacement with improved safety. They can also be used
CC in the development and production of vaccines and biochemical reagents.
CC The present sequence is a corrected version of pol protein
CC from wild type HIV gagpol sequence.
CC The pol coding sequence is codon optimised to be used in the packaging
CC cell line. Codon optimisation results in improved expression of the
CC gagpol protein and reduces the risk of recombination between the transfer
CC vector and gagpol mRNA.
XX Sequence 1003 AA;
```

```
Query Match 57.2%; Score 3117; DB 21; Length 1003;
Best Local Similarity 56.9%; Pred. No. 1.9e-241;
Matches 570; Conservative 178; Mismatches 244; Indels 10; Gaps 6;
QY 32 PSGSSSGSTGEIYAAREKTERAERTIOGSDRGLTAPRAGGDTIOGATNRGLAAPQFSLW 91
Db 9 PQGKAREFSQOTRANSPTREELQ--VWGRDNN-SLSEAGADR-QGTVS--FSPFOITLW 62
QY 92 KRPVVTAIEQOPVEVLDDTGADDSIVAGIELGNNYSPKIVGGIGGFINTKEYKNVEIEV 151
Db 63 QRLVPTIKIGQLKEALLDTGADDTVLEEMNLPGRWKPKMTGGIGGTFKVRQYDQILTEI 122
QY 152 LNKVRATIMTGDPTINIFGRNLTALGMSLNLPAKVPKIMLKPCKDGPKLQWPLT 211
Db 123 CGHKAIGTVLVGTPVNIIGRNLTLQIGCTLNFPISPIETVPVKLKPGMDGPKVQWPLT 182
QY 212 KEKIEALKEICEKMEKEGQLEAEPTNPNTPTFAIKKDKKWKRMILDFRELNKVTQDF 271
Db 183 EEKIKALVEICTEMEKEGKISKIGCPENPTPVFAIKKDKSTKWRKLVDFRELNKRTQDF 242
QY 272 TEIQLGIPHPAGLAKKRITVLDVGDYFSLPHEDRPPYTAFTLPVNNAEPCGRYIK 331
Db 243 WEVOLGIPHPAGLAKKRITVLDVGDYFSLPHEDRPPYTAFTLPVNNAEPCGRYIK 302
QY 332 VLPGWKGSPALFQHTMRQVLEPPRKANKDVIITQYMDLIIASDRTDLEHDRVVLQKLK 391
Db 303 VLPGWKGSPALFQHTMRQVLEPPRKANKDVIITQYMDLIIASDRTDLEHDRVVLQKLK 362
QY 392 LLNGLGFSTPDEKFKQKPPHYHMGYELWPTKWKLOKIQLPQKEITWVNDIQKLVGLNWA 451
```



```

Db 543 KFKLPQKETWEAWWTEYQWATWPEWFEVNTPLVLKWLQLEKEPIIGATFFYVDGAAN 602
Qy 631 ROSKEGAGYVTRDGRKDKVKKLEOTTNQOALEAFAMALTDSPKVINIIVDSQYVMGISA 690
Db 603 RETKLGRAGYVTRDGRQKVVPLTDTTNQKTELQAIHLALQDSGLEVINIVTDSQYALGIQ 662
Qy 691 SQPTSESKIVNQIIEEMIKKEALYVAVPAHKGIGGNOEVDHLVSOGIRQVLEKIEP 750
Db 663 AQDPKSESELVSQIIEQLIKKEKYLAWVPAHKGIGGNEQVDKLVSAIRKVLFDGIDK 722
Qy 751 AQEEHEKHSNWKLSHKFGIPNLVARIIVNSCAQCOQKGEAIIHQVNAELGTWQMDCTH 810
Db 723 AQEEHEKHSNWRAMASDENLPPVVAKEIVASCDKCLKGEAMHGQVDCSPGIWQLDCTH 782
Qy 811 LEGKIIIVAVHVASGFTEAEVIPAESGRQATFALLKLASRWPITHLTDNGANTSEVVK 870
Db 783 LEGKIIIVAVHVASGYTEAEVIPAETQETAYFLLKLGRWPKVTVHTDNGSNFTSTTVK 842
Qy 871 MVAWWIGIEQSFQVYPNPQSQGVVEAMNHHLKNOISRIREQANTIEIVLMAIHCMMFKR 930
Db 843 AACWAGIKQEFPIPNPQSQGVIESMKNELKKIIGQVRDQAEHLKTAQVMAVFIHNFKR 902
Qy 931 RGGIGDMTPSERLINMITTQEIQFLOAKNSKLKDFRVYFREGDQLWKGPGELLWKGE 990
Db 903 KGGIGGYSAGERIIVDIATDIQTKELQKITKIQNFRVYRDSRDPVWKGPAKLLWKGE 962
Qy 991 AVLKVGKTDKIIPRRKAKIIRYDGGQEMDSGSHLEGARE 1032
Db 963 AVVIQNSDIKVVPRKAKIIRDYGKOM---AGDDCVASRQD 1001

```

RESULT 28

```

AAB69284
ID AAB69284 standard; Protein; 1003 AA.
XX AC AAB69284;
XX DT 20-APR-2001 (first entry)
XX DE HIV-1 non-subtype B clone 93BR029-4 pol protein.
XX KW HIV-1; human immunodeficiency virus; non-subtype B; gag; pol; env;
XX KW vpu; vif; vpr; tat; rev; nef; vaccine.
XX OS Human immunodeficiency virus type 1.
XX PN WO200026416-A1.
XX PD 11-MAY-2000.
XX PF 25-OCT-1999; 99WO-US24837.
XX PR 02-NOV-1998; 98US-0184418.
XX PA (UABR-) UAB RES FOUND.
XX PI Hahn BH, Shaw GM, Gao F;
XX DR WPI; 2000-365651/31.
XX PT Novel genomic nucleic acids of non-subtype B human immunodeficiency
XX PT virus type 1 useful for detecting and treating AIDS comprises a
XX PT specific nucleotide sequence -
XX PS Claim 41; Fig 15; 131pp; English.
XX CC The present invention provides the protein and coding sequences for a
XX CC number of human immunodeficiency virus (HIV) type 1 non-subtype B
XX CC isolates. The sequences shown include the near full-length coding
XX CC sequences from each isolate, and the env, pol, vif, vpr, vpu, gag, tat,
XX CC rev and nef proteins. These can be used to detect the presence of HIV-1
XX CC in a sample and to produce antibodies against non-subtype B HIV-1 virus.

```

These antibodies can be used in vaccines to prevent and treat HIV infection.

CC Sequence 1003 AA;

Query Match 57.2%; Score 3116.5; DB 21; Length 1003;
Best Local Similarity 56.9%; Pred. No. 2e-241;
Matches 570; Conservative 173; Mismatches 250; Indels 9; Gaps 4;

```

Qy 32 PSGSSSGSTGIYAAAREKTERAERETIQGSDRGUTAPRAGGDTTQAGATNRGLAAPQFSLW 91
Db 9 PQGKAREFPSSQTRANSPTS---ELQVWGRGNSLSSETGADRQGDVSFGF--PQITLW 62
Qy 92 KRPPVVTAYIEQPPVEVLDDTCADDSIVAGIELGNNSPKIVGGTGGFTNTREYKNVELEV 151
Db 63 QRLVTVKIGQLKEALDDTGADDTVLEENLPGRWKPMGIGGFKVQYQDQIPIEI 122
Qy 152 LNKVRATITMGDTPINIFGRNLTALGMSLNLPVAKVEPIKIMLPGKDGPKLRQWPLT 211
Db 123 CGRKATGVLVGPVPIVNIIGRNLLTQIGCTLNFPIETVPIVVKLPGMDGPRVKQWPLT 182
Qy 212 KEKEALKEICEKEKEGQLEEAAPTNPYNTPTFAIKKKKNKRWMLIDFRELKNKVODF 271
Db 183 EEKIKALTEICTEMEKEGKISKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELKRTQDF 242
Qy 272 TEIOLGIPHPAGLAKKRITVLDVGDVAFVSTPLHEDPRPYTAFTLPSVNNAPCGKRYIK 331
Db 243 WEVOLGIPHPAGLAKKRSVTVLDVGDVAFVPLDOKFRKYTAFTPTNNPTTGLURQYN 302
Qy 332 VLPQGWKSPAIQHTMRQVLEPFERKANKDVIILQYMDDIILASDRTDLEHDRVVLQKE 391
Db 303 VLPQGWKSPAIFOSSMTKILEPFRKQNPDIYIYQMDLVGSDLEIGHQRTKEELRQ 362
Qy 392 LLNGLGFSTPDEKQDPPYHWMGYELMPTTKWLQKIQLPQKEIWTWVNDIQKLVGLNWA 451
Db 363 HLLRWGFTTDPKHKQKPPFELWNGYELHPDKWTQPIVLPEKDSWTWVNDIQKLVGLNWA 422
Qy 452 AOLYPGIKTHLCRLIRGKMTLTEEVQWTELAELLENRIILSQEQGHGYEKEKELEA 511
Db 423 SQIYAGIKVROLCKLKGTKALTEVPLTABAELAELENRELLPEVHGVIYDSDKDLIA 482
Qy 512 TVQKDQENQWTKYKIHQEE-KILKVGKYAKVKNTHNTNGIRLLAAQVQVQKIGKALVIGRIP 570
Db 483 EIQKQGGQWYQIYQEPYKNLKTKGYARMGAHTNDVKQLTEAVQKITTESIIVGWKIP 542
Qy 571 KFHLPVREIWEQWMDNVTWIPDWDVFTPLVRLAFNLVGDPIPGATFTYDSCN 630
Db 543 KFKLPQKETWEAWWTEYQWATWPEWFEVNTPLVLKWLQLEKEPIVGAETFFYVDGAAN 602
Qy 631 ROSKEGAGYVTRDGRKDKVKKLEOTTNQOALEAFAMALTDSPKVINIIVDSQYVMGISA 690
Db 603 RETKLGRAGYVTRDGRQKVVPLTDTTNQKTELQAIHLALQDSGLEVINIVTDSQYALGIQ 662
Qy 691 SQPTSESKIVNQIIEEMIKKEALYVAVPAHKGIGGNOEVDHLVSOGIRQVLEKIEP 750
Db 663 AQDPKSESELVSQIIEQLIKKEKYLAWVPAHKGIGGNEQVDKLVSAIRKVLFDGIDK 722
Qy 751 AQEEHEKHSNWKLSHKFGIPNLVARIIVNSCAQCOQKGEAIIHQVNAELGTWQMDCTH 810
Db 723 AQEEHEKHSNWRAMASDENLPPVVAKEIVASCDKCLKGEAMHGQVDCSPGIWQLDCTH 782
Qy 811 LEGKIIIVAVHVASGFTEAEVIPAESGRQATFALLKLASRWPITHLTDNGANTSEVVK 870
Db 783 LEGKIIIVAVHVASGYTEAEVIPAETQETAYFLLKLGRWPKVTVHTDNGSNFTSTTVK 842
Qy 871 MVAWWIGIEQSFQVYPNPQSQGVVEAMNHHLKNOISRIREQANTIEIVLMAIHCMMFKR 930
Db 843 AACWAGIKQEFPIPNPQSQGVIESMKNELKKIIGQVRDQAEHLKTAQVMAVFIHNFKR 902
Qy 931 RGGIGDMTPSERLINMITTQEIQFLOAKNSKLKDFRVYFREGDQLWKGPGELLWKGE 990
Db 903 KGGIGGYSAGERIIVDIATDIQTKELQKITKIQNFRVYRDSRDPVWKGPAKLLWKGE 962

```

```

QY 991 AVLKVGTDIKIIPRKAIIIRDYGGQEMDSGSHLEGARE 1032
Db 963 AVVIQDNDIKVVPKRKVIIRDY--KQAGGDCVAGRQDE 1002

RESULT 29
AAR29705
ID AAR29705 standard; Protein; 1003 AA.
XX AAR29705;
AC
XX
XX
DT 25-MAR-2003 (updated)
DT 26-APR-1993 (first entry)
XX
XX
DE pol gene decoded from viral DNA from HTLR.
XX
KW express proteins; lymphadenopathy syndrome; AIDS; HIV; HTLV;
KW serological immunoassays; antibodies to HTLR; monoclonal antibodies;
KW probes; ss.
XX
OS Human T cell lymphotropic retrovirus.
XX
XX
XX
PN EP518443-A1.
XX
XX
PD 16-DEC-1992.
XX
XX
PF 30-OCT-1985; 92EP-0201711.
XX
XX
PR 31-OCT-1984; 84US-0667501.
PR 30-JAN-1985; 85US-0696534.
PR 06-SEP-1985; 85US-0773447.
XX
XX
PA (CHIR ) CHIRON CORP.
XX
XX
PI Barr PJ, Dina D, George-Nascimento C, Hallewell R;
PI Luciw PA, Parkes D, Pescador RS, Steimer K, Truett M;
XX
XX
DR WPI; 1992-417329/51.
DR N-PSDB; AAQ31936.
XX
XX
PT Recombinant DNA construct including replication system recognised
PT by unicellular microorganism - used to form recombinant proteins
PT for diagnosing AIDS and lymphadenopathy syndrome
XX
XX
PS Example 5; Fig 2; 32pp; English.
XX
XX
CC This sequence was decoded from the pol gene from proviral HTLR DNA.
CC Proteins associated with lymphadenopathy syndrome and/or AIDS may
CC be used in serological immunoassays to detect antibodies to HTLR.
CC The polypeptides can be used alone or in fusion constructs to
CC produce antisera or monoclonal antibodies which may be used for
CC therapy or diagnosis.
CC (Updated on 25-MAR-2003 to correct PN field.)
CC (Updated on 25-MAR-2003 to correct PF field.)
XX
XX
SQ Sequence 1003 AA;

Query Match 57.2%; Score 3116; DB 13; Length 1003;
Best Local Similarity 56.7%; Pred No. 2.2e-241;
Matches 566; Conservative 180; Mismatches 241; Indels 12; Gaps 6;

QY 38 GSTGEIYAAREKTERARETIQ---GSDRGLTAPRAGGDTIQGATNRGLAAPQFSLWKR 94
Db 11 GKAREFSSQTRANSPTRELQVWGGENSL--EAGADR-QGTVSFNF--PQITLWQRP 65
QY 95 VVTAYIEGQPVVLLDTGADDSIVAGIELGNYSKPIVGGIGGFINTKEYKNVEVLNK 154
Db 66 LVTIRIGGLKCALLDTGADDTVLEEMNLPGKWKPMIGGIGGFIKVQDQIPVEICGH 125
QY 155 KVRATIMTGDPTPINFGNLTALGMSLNPVAKVEPIKIMLPGKGDGPKLRQWPLTKRK 214
Db 126 KAIGFVLVGPFPVNIIGNLLTQIGCTLNFPISPIETVPVKLPGMDGPKVQWPLTEK 185

```

```

QY 215 IEALKEICEKMEKQLEBEAPPTNPTNPTFAIKKKKKNKRMILDFRELNKVTDQFTEI 274
Db 186 IKALVEICTEMEKEGKISKIGPENPTNPTNPTFAIKKKKKNKRMILDFRELNRTOFWFWE 245
QY 275 QLGIPHPAGLAKRRITVLDVGDAYFSIPLHEDFPYTAFTLPSVNNAPGKRYIYKVL 334
Db 246 QLGIPHPAGLAKKKSIVLDVGDAYFSVPLDKDFRYTAFTLPSINNTPGIRYQYNVLP 305
QY 335 QGKGSIPAIFQHTMRQVLEPFRKANKDVIIQYMDLILASDRTDLHDHVVVLLQKLELN 394
Db 306 QGKGSIPAIFQSSMTKILEPFRKQNPDIYQYMDLIVGSDLEIGHQHRKTEELRQHLL 365
QY 395 GLGFSTPDEKFOKDPYPHMGYELPFTKWKLOKIQLPQKEIWTVDNDIQKLVGLVNWAAOL 454
Db 366 RWGFTTPDKKHQKEPPFLMMGYELHPDKWTQVQIMLPKXDSWTVNDIQKLVGLVNWASQI 425
QY 455 YGSIKTKHLCLIRGKMTLITEEVOMTELAEALEENRIILSQBOEGHYVQEEKELPATVQ 514
Db 426 YAGIKVKQLCKLLRGTKALTEVIPLEEAELEENREILKEPVHEVYDPSKDLVAEIQ 485
QY 515 KDQENQWYKIHQEE--KILKVGKYAKVKNTHNTGIRLLAQVVKIGKEALVIGRIPKPH 573
Db 486 KQGQGWYQIYQEPFKNLKTCYARMCAHNDVYKQTEAVQKYSTESIVLWIKPIPKFK 545
QY 574 LPVEREINBQWMDNYQWTVIPDWDVFSTPPVLRLAFNLVGDPIPAEATFYTDGSCNROS 633
Db 546 LPIQKETWEAMWMEYQATWIPWEPFVTPPLVKLWYQLEKEPIVGAETFYVDGAANRET 605
QY 634 KEGKAGYVTDGKDKVKLEQTTNQOAELEAFAMALTDGPKVNIIVDSQYVNGISASOP 693
Db 606 KLGKAGYVTDGKGRQVWSTADTTNQTQLQAIHLALQDQSGLEVINVTDSQYALGIIQAO 665
QY 694 TESKIVNQIIEEMIKKEAIYVAVWPAHKGIGGQVNDHLVSQGRVLFLEKLEPAQE 753
Db 666 DASESELYSQIIEQLIKKEKYLAWVPAHKGIGGNEQVDKLVKSAGIRKVLFLNGIDKAE 725
QY 754 EHEKYHSNVELSHKFGIPNLVARQIVNSCAQCKQGEAIHQVNAELGTQWMDCTHLEG 813
Db 726 EHEKYHSNWRAMASDFNLPPVVAKEIVASCDQLKGEAMHGQVDCSPGIWOLDCTHLEG 785
QY 814 KIIIVAVHVSAGFIEAEVTPQESGRQTAFLFLKLASRWPITHLHTDNGANFTSQEVKVA 873
Db 786 KIILVAVHVSAGYIEAEVTPAQGTQETAYFLKLGRMPVKTIHTDNGSNFTSTTVKAC 845
QY 874 WMIGIEQSFQVPYNPQSQGVVEAMNHLKNQISRIREQANTLETIVLMAIHCNFKRRGG 933
Db 846 WWAGIKQFEGIPYNPQSQGVVESMNNELKKIIGVYRQDAEHLKTAVQMAVTHNFKRGG 905
QY 934 IGDWTPSERLINMITTEOEIOPLOAKNSKLDKDFRYVFRGRDQLWKGPGELLWKGAVL 993
Db 906 IGGYSAGERIVDIIATDITQKELQKITKIQNFYVYRDNDKPLWKGPAKLWKGAVV 965
QY 994 VKVGTDTKITIPRKAIIIRDYGGQEMDSGSHLEGARE 1032
Db 966 IQDNDIKVVPKRKAIIIRDYGGQEMDSGSHLEGARE 1001

RESULT 30
AAR70861
ID AAR70861 standard; Protein; 1003 AA.
XX
XX
AC AAR70861;
XX
XX
DT 25-MAR-2003 (updated)
DT 23-APR-1991 (first entry)
XX
XX
DE Sequence encoded by LAV/HTLV III pol gene.
KW Glycoprotein gp 110; gp 41; AIDS vaccine; diagnosis.
XX
XX
OS LAV/HTLV III.
XX
PN W08702038-A.

```


CC A fragment of DNA approximately 200-500 base pairs in length is
CC ligated into a vector and used to transform E.coli. These cells then
CC express a polypeptide which is immunoreactive with HTLV-III-specific
CC antibody. The HTLV-III polypeptides can be used for the production
CC of antibodies, in immunosays for the detection of HTLV-
CC III-specific antibodies and in vaccines for the prevention of AIDS.
CC The antibodies can also be used to detect HTLV-III polypeptides.
CC (Updated on 25-MAR-2003 to correct PN field.)
CC (Updated on 25-MAR-2003 to correct PF field.)
CC (Updated on 25-MAR-2003 to correct PR field.)
XX
SQ Sequence 1015 AA;

Query Match 57.1%; Score 3114; DB 14; Length 1015;
Best Local Similarity 56.4%; Pred. No. 3.3e-241;
Matches 566; Conservative 177; Mismatches 241; Indels 20; Gaps 4;

QY 46 ARE---KTERAERTIGSDRGITAPRAGDGTIQGATNRGLAA-----PQFS 89
Db 13 AREFSSEQTRANSPTISSEQTRANSPTRELQVWGRDNNSLSEAGADROGTVSFNFPQIT 72

QY 90 LAKRPVVTAYIEGQPEVLLDTCADDSIVAGIELGNYSKPIVGGIGGINTKEYNVEI 149
Db 73 LMQRPLVTIKIGQLKEALLDTCADDTVLEEMSLPGRWKPKMIGGIGGIFKVRQYDQILI 132

QY 150 EVLNKKVRATIMTGTDPINIFGRNLTALGMSLNLPAKVEPIKIMLKPGKDGPKLRQWP 209
Db 133 EICGHAIGTVLGPPTVNLIGRNLLTGCTLNFPISPIETVPVKLGKMGDPKVKQWP 192

QY 210 LTKEIEALKEICEKMEKQGLDEEAPNTPTFAIKKKKKNKWRMLIDFRELNKVTQ 269
Db 193 LTEEIKALVIEITEMEKESKIGPENPNTPVFAIKKKKSTKRWKLVDFRELNRRTQ 252

QY 270 DFEIOLGPHAPLAGLAKRITVLDVGDVAFESPLHEDRPYTAFTLPSVNAEPCKRYI 329
Db 253 DFWEVLGPHAPLAGLAKKSVTVLDVGDVAFVPLDEDFKRYTAFTIPSINNETPGSGQ 312

QY 330 KYVLPGQWKGSPAIFQTRQVLEPPFRKANKDVIITQYMDLILIASDRTDLEHDRVVLQ 389
Db 313 YNVLPGQWKGSPAIFQSSMTKILEPFRKQNPDIIVQYMDLVGSDLEIGQHRKIEEL 372

QY 390 KELLNGLSTPDEKQKQPPHMGYELWPKWKLOLQLOKELTWTVDNQKLVGNL 449
Db 373 ROHLRWGFTTPDKKQKQPPFLMWGYELHPDKWTQIPLPKDSWTVDNQKLVGNL 432

QY 450 WAAOLYPGKTKHLCLIRGKMTLVEEQWTELAELEENRILISQEQEGHYQEKEKEL 509
Db 433 WASQIYPGIKVROLCKLRGKTALTEVILTEEALELEENRILKEPVGHYVDPSKDL 492

QY 510 EATVQKQENQWYTKIQHEE-KILKGYKAKVKNTHNGIRLLAQVQVQKIGKEALYIWR 568
Db 493 IAEIQOGQGWYQIYQEPFKNLTKGYARMGAHTNDVKQLTEAVQKITTESIYIWK 552

QY 569 IPKFHLPLVERIEWEQWHDNYQVWTIPDNDVSTPPLVRLAFNLVGDPIPGATFTDGS 628
Db 553 TPKEFLPIQKETWETWTETWQATWIPEWFEVNTPLVWLWYOLEKEPIVGATFTYVDGA 612

QY 629 CNRQSKRAGYVTDGKDKVKLEOTNQOAELEAFAMALDSDGPKVNIIVDSQXVMGI 688
Db 613 ASRETKLGRAGYVNRGRQVYLTHTTNQKTELQAINLALQDSGLEVNIIVDSQYALGI 672

QY 689 SASQPTSESKIVNQIIEEMIKKEAYIYVAVWPAHKGIGGQNEVDHLSOGIRQVLFLEKI 748
Db 673 IQAQPDKSESELVNIQIEILIKKEVYLAWVPAHKGIGGQNEVDKLVSAIGIRKILFDGI 732

QY 749 EPAQEEHEKXVNSKLSHFKGIPNLVARIQVNSCAQCOOKGEAIGHQVNAELGTWQMC 808
Db 733 DKAQEEHEKXVNSRAMSDFNLPPVVAKEIYVASCDCQLKGEAMHGQVDCSPGIIQLDC 792

QY 809 THLEGRITIVAVHVASGFTAEAVIPOESGRQTFALFLLKLSRWPIPTHLTNDGANTSOE 868
Db 793 THLEGRVILVAVHVASGYEAEVIPAETQETAYFLLKLAGRPVPTVTHTNDGNSNFTSAT 852

QY 869 VKVMAWVIGIEQSFQVYPNPQSQGVVEAMNHHLKNOISRIREQANTETIVLMAIHCNMF 928
Db 853 VRAACWAGIKQEFGIPYNPQSQGVVSNKELKKIIGVQDAEHLKTAVQMAVEITHNF 912

QY 929 KRRGGIGDMTPSRRLNMTTTEQEIQFLOAKNSKLDKFRVYFREGRDQLWKGCELLWKG 988
Db 913 KRRGGIGGYSAGERIVDIATDIQTKELQKITKIONFRVYRDSRNLWKGPAKLLWKG 972

QY 989 EGAVLVKVGTDKIIIPRRKAKIIRDYGGGRQEMDSSGSHLEGARD 1032
Db 973 EGAVVIQDSDIKVPPRRKAKIIRDYGKQ---AGDDCVASROD 1013

RESULT 32
AAP61507
ID AAP61507 standard; Protein; 1012 AA.
XX AAP61507;
XX XX
DT 08-JUN-1991 (first entry)
XX Sequence of reverse transcription ("pol").
XX HIV; LAV; AIDS; diagnosis; vaccine.
XX HTLV-IIIIB/H9 cells (ATCC CRL 8543).
XX EP187041-A.
XX 09-JUL-1986.
XX 23-DEC-1985; 85EP-0309454.
PR 24-DEC-1984; 84US-0685272.
PR 04-DEC-1985; 85US-0805069.
XX (GETH) GENENTECH INC.
XX Capon DJ, Lasky LA;
XX WPI; 1986-177602/28.
DR N-PSDB; AAN60288.
XX Acquired immune deficiency syndrome polypeptide(s) - obt'd. by
PT molecular cloning etc. and used for diagnosis and in vaccines
PT against virus disease
XX Example; fig 2; 125pp; English.
XX A comparison of AAN60287 with the cDNA of the HTLV-III genome
CC revealed one particular clone, designated p7.11 which contained a
CC DNA sequence encoding this peptide (AAP60308) sequence. This approx.
CC 2.2 kilobase covers the precursor gag region and encodes, 5' to 3',
CC p-12, p-15, p-24 a second p-15 protein, and approx. 300 extra base
CC pairs 3' to the gag region (see AAN60288).
SQ Sequence 1012 AA;

Query Match 57.1%; Score 3113; DB 7; Length 1012;
Best Local Similarity 56.2%; Pred. No. 4e-241;
Matches 563; Conservative 183; Mismatches 249; Indels 6; Gaps 3;

QY 33 SGSSSGSTGEIYAAAREKTERAERTIQGSDRGITAPRAGDGTIQGATNRGLAAPQFSLWK 92
Db 15 SEQTRANSPTISSEQTRANSPTRELQVWGRDNNSPSEAGADROGTVSFNF--PQITLMQ 72

QY 93 RPVVVTAYIEGQPEVLLDTCADDSIVAGIELGNYSKPIVGGIGGINTKEYNVEI 152
Db 73 RPLVTIKIGQLKEALLDTCADDTVLEEMSLPGRWKPKMIGGIGGIFKVRQYDQILTEIC 132

QY 153 NKKVRATIMTGTDPINIFGRNLTALGMSLNLPAKVEPIKIMLKPGKDGPKLRQWPLTK 212
Db 133 GHKAIGTVLGPPTVNLIGRNLLTGCTLNFPISPIETVPVKLGKMGDPKVKQWPLTE 192

[illegible]

Db 903 KGGIGYSAGERIVDIIATDIOTKELQKQITKIONFRVYVYRDSRDPVWVKPAKLLWKGE 962
 QY 991 AVLKVGVDIKIIPRKAIIIRDYGGROEMDSGSHLEGARE 1032
 Db 963 AVTIQNSDIKVPFRKAIIIRDYQKQ---AGDDCVASRQD 1001

RESULT 35
 ID AAY70600 standard; Protein; 1003 AA.
 XX AAY70600;
 AC AAY70600;
 XX 18-JUL-2000 (first entry)
 DT Wild type Human immunodeficiency virus pol protein.
 DE HIV; gag; packaging cell line; lentivirus; retroviral vector particle;
 KW pol; gene therapy; gene replacement; vaccine; biochemical reagent.
 KW Human immunodeficiency virus.
 OS
 XX WO200015819-A1.
 PN 23-MAR-2000.
 PD 10-SEP-1999; 99WO-US20675.
 PF 11-SEP-1998; 98US-0100022.
 PR 12-SEP-1998; 98US-0100063.
 XX (CHIL-) CHILDRENS MEDICAL CENT.
 PA Gray JT, Mulligan RC;
 PI WPI; 2000-271455/23.
 XX N-PSDB; AAZ52052.
 DR New packaging cell line for producing a viral accessory protein
 PT independent HIV derived retroviral vector particles, useful in gene
 PT therapy or gene replacement -
 XX Disclosure; Fig 9; 62pp; English.

The patent discloses new packaging cell line for producing a viral accessory protein independent lentivirus, preferably human immunodeficiency virus (HIV), derived retroviral vector particles. The packaging cell line comprises a mammalian cell, a retroviral DNA comprising a coding sequence for a lentivirus, preferably HIV, gagpol, where the coding sequence has been mutagenised to improve expression of the viral gagpol proteins, a second retroviral nucleotide sequence comprising the coding sequence for a heterologous envelope protein and a third retroviral nucleotide sequence comprising a DNA sequence of interest and lentivirus, preferably HIV, cis-acting sequences required for packaging, reverse transcription and integration. The packaging cell lines and viral particles can be used for gene therapy or gene replacement with improved safety. They can also be used in the development and production of vaccines and biochemical reagents. The present sequence is a pol protein of wild type HIV gagpol. The pol coding sequence is codon optimised to be used in the packaging cell line. Codon optimisation results in improved expression of the gagpol protein and reduces the risk of recombination between the transfer vector and gagpol mRNA.

XX Sequence 1003 AA;
 SQ

Query Match 57.1%; Score 3112; DB 21; Length 1003;
 Best Local Similarity 56.9%; Pred. No. 4.7e-241;
 Matches 570; Conservative 177; Mismatches 245; Indels 10; Gaps 6;

QY 32 PSGSSSGTGEIYAAREKTERAERTTQSGDRGTTAPRAGGDTTQGNRGLAAPQPSLW 91
 Db 9 PQGKAREFSSQTRANSPTRELO--VWGRDNN-SLSEAGADR-QGTVS--FSFPQITLW 62

QY 92 KRPVVYATIEGQPVVLLDTGADDSIVAGIELGNNYSKPIVGIGGFINTKYEKVEIEV 151
 Db 63 QRPLVTIKIGOLKEALLDTGADDTVLEEMNLPGRWKPKMIGGIGGFVKVGYDQILIEI 122
 QY 152 LNKVVRATMTGDTPIINIFGRNLPALGMSLMLPVAKVEPKIMILKPKDGDGPKLQWPLT 211
 Db 123 CGHKAIGTVLVGPTVNIIGRNLITQIGCTLNFPIISPIETVPVKLPGMDGPKVQWPLT 182
 QY 212 KEKTEALKEICEKMEKEGLEEAPPTNPYNTTEALKKKDKNKRMLDFRELNLVTODF 271
 Db 183 EEKIKALVEICTEMEKEGKISKIGPENPINTPVFAKKKDSWKRLKLVDFRELNRTOOF 242
 QY 272 TEIQLGIPHPAGLAKKRITVLDVGDAYFSIPLEDHFRPYTAFTPLSPVNNAPGKRYIK 331
 Db 243 WEVQLGIPHPAGLKOKSVTVLDVGDAYFSVPLDKDFRYTAFTPIPSINNETPGIRYQYN 302
 QY 332 VLPQGWKSPAIFQHTMRQVLEPPFRKANKDVIIIOYMDILLIASORTLEHDRVVLQKE 391
 Db 303 VLPQGWKSPAIFQCSMTKILEPPFRKQNPDIYQYMDLLYVGSDELIGQHRTKIELRQ 362
 QY 392 LLNGLGFSTPDEKFOKDPYHWMGYELMPTTKWKLOLPOKEIWTVNDIOKLVLGNAWA 451
 Db 363 HLLRWGFTTPDKKHQKEPFLWMGYELHPDKWTQVPIVLPEKDSWTVNDIOKLVLGNAWA 422
 QY 452 AQLYPGIKTKHLRLIRGKMTLITEVQWTEALAELEENRIILSOEGHYHQQEKELEA 511
 Db 423 SQIYAGIKVRQICKLLRGTKALTEVVPLEAELEAENREILKEPVHGVYDPSKDLIA 482
 QY 512 TVQKQDNQWYKIHQEE-KILKVGKYAKVKNHTNGIRLLAQVYVOKIGKEALVIGRIP 570
 Db 483 EIQOGQGWYQIYQEPFKLTKGYARMKAGTNDVKQLTEAVQKATIESIVIGKTP 542
 QY 571 KEHLIPVEREINQWWDNTWQVTPIDWDFVSTPPLVRLAFNLVGDPIPGAETFFYDGSN 630
 Db 543 KEKLPQIKETWEAMWTEYQATWIPWEFEVNTPLVKLWYQLEKEPIIGAETFFYVDGAAN 602
 QY 631 QSGREGKAGYVTRGKDKVKKLEQTTNOQAAELAEAFAMALTDSGPKNVIIVDSQYVNGISA 690
 Db 603 RETKLKAGYVTRGSRQKVPVPLTDITNQKTELQAIHLALQDSGLEVNIVTDSQYALGIQ 662
 QY 691 SPTSESKIVNQIIEEMKKEAIVAVWPAHKGIGNQEVNHLVSGIRQVLFLEKLEP 750
 Db 663 AQPKSESELVSQIIEQLIKKEKYLAWVPAHKGIGNEQVDGLVSAGIRKVLFDGDK 722
 QY 751 AQEEHEKYSNVKLSHKFGIPNLVAROIVNSCAOQKGEAIGHGVNAELGTWQMDCTH 810
 Db 723 AQEEHEKYSNWRAMASDFNPPVVAKEIVASCDCOLKGEAMHGVQDCSPGIWQDCTH 782
 QY 811 LEGKIIIVAVHVASGFIEAEVTPQESGRQTALFLKLASRWPITHLHTDNGANFTSOEVK 870
 Db 783 LEGKVLVAVHVASGYIEAEVTPAQGTAYFLLKLGRWPVKVTHVTDNGSNFTSTTVK 842
 QY 871 MVANWIGTEQSGFVYPNPOSQGVFAMNHHLKQISRIREQANTTETIVLMAIHCMPKR 930
 Db 843 AACWAGIKQFEGIPNPQSQGVIESMKNLKKIIGQVRDQAEHLKTAQMAVFLHNEKR 902
 QY 931 RGGIDMTSPERLINMITTEQEIQLOAKNSKLDFRVYFREGRDOLAKGPELAKWKEG 990
 Db 903 KGGIGYSAGERIVDIIATDIOTKELQKQITKIONFRVYVYRDSRDPVWVKPAKLLWKGE 962
 QY 991 AVLKVGVDIKIIPRKAIIIRDYGGROEMDSGSHLEGARE 1032
 Db 963 AVTIQNSDIKVPFRKAIIIRDYQKQ---AGDDCVASRQD 1001

RESULT 36
 AAP60347
 ID AAP60347 standard; protein; 1015 AA.
 XX AAP60347;
 XX 25-MAR-2003 (updated)
 DT

Db 631 LGCTLNFPISPIETVPVVKLPGMGPKVKQWPLTEEKIKALTAICDEMEKEGKITKIGPE 690
 QY 238 NPYNTPTFAIKKKDKKRWMLIDFRELANKVTQDETEIOGLGPHAGLAKKRITVLDVGD 297
 Db 691 NPYNTPTFAIKKKDKKRWMLIDFRELANKVTQDETEIOGLGPHAGLAKKRITVLDVGD 750
 QY 298 AYSFISPLHEDFRPYTAFTPLPSVNNAPGKRIYKVLPGQWKGSPAIFQTHMRQVLEPRFK 357
 Db 751 AYSFISPLYEDFRPYTAFTIPSRNNETPGIRYQYNNVLPQWKGSLAIFQSSMTKLEPRFK 810
 QY 358 ANKDVIIIOYMDLIIASDRDLEHDRVVLQKELLNGLGFSTPDEKFKQKPPYHMGYE 417
 Db 811 QNPGIYIQQYMDLIIASDRDLEHDRVVLQKELLNGLGFSTPDEKFKQKPPYHMGYE 869
 QY 418 LWPWKWKLOKQLPKOKETWVNDIOKLVGNLWAAQLYPGIKTHLCLRLIRGKMTLTREV 477
 Db 870 LHPDKWTVPQTLPEKDSWTVNDQKLVGNLWAAQLYPGIKTHLCLRLIRGKMTLTREV 929
 QY 478 QWTELAELAELENRIILSQEGHYQBEKELEATVQKQDNQWYTKIHQEE-KILKVGK 536
 Db 930 PLTEAELELAENREILKEPVHGVYDPSKDLIAEIQKOGQEQWYQIYQEPFKNLTKG 989
 QY 537 YAKVKNTHNGIRLLAQVQKIGKEALVINGRIPKPHLPVEREIQWQWMDNYQWTPD 596
 Db 990 YAKWRTAHTNDVKQLEATEAVQKIGKEALVINGRIPKPHLPVEREIQWQWMDNYQWTPD 1049
 QY 597 WDFVSTPLPLVRLAFNLVGDPIPGAETFYTDGSCNRSGKAGYVTDGKDKVKKLEOTT 656
 Db 1050 WEFVNTPLVRLAFNLVGDPIPGAETFYTDGSCNRSGKAGYVTDGKDKVKKLEOTT 1109
 QY 657 NQAELEAFAMALTDGPKVNIIVDSQYVMGISAQPTSESKIVNQIIEEMIKKEATYV 716
 Db 1110 NQAELEAFAMALTDGPKVNIIVDSQYVMGISAQPTSESKIVNQIIEEMIKKEATYV 1169
 QY 717 AWVPAHKGIGQNEQVHLVSGIRQVLFLEKIEPAQEEHEKYSNVKLSHKFGIPNLVA 776
 Db 1170 SWVPAHKGIGQNEQVHLVSGIRQVLFLEKIEPAQEEHEKYSNVKLSHKFGIPNLVA 1229
 QY 777 RQIVNSCAQCKQGEAIGHQVNAELGTWQMDCTHLEGKIIIVAVHVASGFTAEVPOES 836
 Db 1230 KEIVASDCOLKGEAMHGVQVDCSPGIWQLDCTHLEGKIIIVAVHVASGFTAEVPOES 1289
 QY 837 GROTALFLKLASRPIITHLTDNGANFTSQEVKMAVWIGIEQSFQYPPNQSQGVVEA 896
 Db 1290 GROTALFLKLASRPIITHLTDNGANFTSQEVKMAVWIGIEQSFQYPPNQSQGVVEA 1349
 QY 897 MNHLKKNQISIRIQANTITIVLMAHGMFKRGGIGDMPSERLNMTTQEQIOFL 956
 Db 1350 MNHLKKNQISIRIQANTITIVLMAHGMFKRGGIGDMPSERLNMTTQEQIOFL 1409
 QY 957 QAKNSKLKDFRVYFREGDQLWKGPELLWKGAGVAVKVGTDIKIIPRRKAKIIRDXG- 1015
 Db 1410 QAKNSKLKDFRVYFREGDQLWKGPELLWKGAGVAVKVGTDIKIIPRRKAKIIRDXG- 1469
 QY 1016 -----GRQEMD 1021
 Db 1470 QMAGADCAGRQEMD 1484

RESULT 40

AA019387
 ID AA019387 standard; Protein: 1003 AA.

XX
 AC AA019387;

XX
 DT 10-DEC-2002 (first entry)

XX
 DE Lymphadenopathy-associated virus pol protein.

XX
 KW Lymphadenopathy-associated virus; LAV; HIV; human immunodeficiency virus;

KW Lymphadenopathy; AIDS; gag; pol; orf Q; env; orf F; anti-HIV; virucide;

XX

OS Lymphadenopathy-associated virus.

XX US6440657-B1.

XX 27-AUG-2002.

XX 06-JAN-2000; 2000US-0478492.

PF 16-NOV-1984; 84GB-0029099.

PR 18-OCT-1985; 85CA-0493377.

PR 03-NOV-1992; 92US-0970954.

PR 17-OCT-1986; 86US-0920119.

PR 01-AUG-1989; 89US-0390499.

PR 06-DEC-1990; 90US-0622278.

PR 20-AUG-1991; 91US-0747506.

PR 01-OCT-1993; 93US-0130565.

PR 06-JUN-1995; 95US-0466920.

PR 05-DEC-1983; 83US-0558109.

PR 28-FEB-1985; 85US-0706562.

PR 30-AUG-1985; 85US-0771230.

PR 30-AUG-1985; 85US-0771247.

PR 30-AUG-1985; 85US-0771248.

XX (INSP) INST PASTEUR.

PA (CNRS) CENT NAT RECH SCI.

XX Montagnier L, Krust B, Chamaret S, Clavel F, Chermann J;

PI Barre-sinoussi F, Alizon M, Sonigo P, Cole S, Danos O;

PI Wain-Hobson S;

XX WPI; 2002-711525/77.

DR N-PSDB; AAL49920.

XX Novel chemically synthesized Human Immunodeficiency Virus envelope

PT protein, useful in vitro diagnostic method for the detection of the

PT presence or absence of antibodies that bind to antigens of HIV-1

PT retrovirus

XX Disclosure; Fig 3; 31pp; English.

PS The present invention relates to a chemically synthesised env peptide of

CC Human immunodeficiency Virus (HIV) of less than 150 amino acid residues.

CC Such a peptide is useful in an in vitro diagnostic method for the

CC detection of the presence or absence of antibodies that bind to antigens

CC of HIV-1 retrovirus, by contacting the peptide with a biological fluid

CC for a time and under conditions sufficient for the peptide and antibody

CC in the biological fluid to form a peptide-antibody complex, detecting the

CC formation of the peptide-antibody complex by comparing the formation of

CC peptide-antibody complex formation with a control sample, where the

CC formation of peptide-antibody complex is correlated with the presence of

CC antibodies that bind to antigens of HIV-1 retrovirus in the biological

CC sample. It can also be used for preparation of vaccine compositions

CC against AIDS or related syndromes. The present sequence is the

CC Lymphadenopathy-associated virus LAV pol protein. The LAV virus was

CC isolated from a patient with AIDS.

XX Sequence 1003 AA;

SQ Query Match 57.1%; Score 3111; DB 23; Length 1003;

Best Local Similarity 56.6%; Pred. NO. 5.7e-241;

Matches 567; Conservative 176; Mismatches 235; Indels 24; Gaps 5;

QY 52 RAERETIQSGDRGLTA--PRAGGDT-----TQATNRGLAA-----PQFSIAW 91

Db 3 REDLAFLOGKAREFSSEQTRANSPTRRRELQVWGRDNNLSBAEADROQTVSFNFQITLW 62

QY 92 KRPVVYATIEGQVPEVLDTGADDSIVAGIELGNNYSKPIVGGIGGFTINTKEYKNVEIEV 151

Db 63 QRPLVTIKIGQLKEALLDTGADDTIVLEMSLPGRWPKMIGGIGGFTIKVQYDILIEI 122

QY 152 LNNKVRATIMTGDPTINIFGRNLTALQMSLNPVAKVEPKIMLPGKDGPKLRQWPLT 211

Db 123 CGHKAIGTVLVGPTPVNIIGNLLTQIGCTLNFSPISPIETVPVVKLPGMGDPKVKQWPLT 182

Db 843 AACWAGIQEFGIPYNQSGVYESMKNELKKIIQIRDOAEBLKTAVQMAVFIHFKR 902

Qy 931 RGIGDMTPSERLINMITTEQETQFLQAKNSKLKDFRVYFREGDRDLWGPGLLWKGE 990

Db 903 KGGIGYSAGERTIDIATDIQRELQKQIKIQRNFRVYRDRDPVWKGPAKLLWKGE 962

Qy 991 AVLKVGTDIKIIPRKAIRIDY-----GROEMD 1021

Db 963 AVVIQDNSEIKVPRKAKIIRIDYGKQMGAGDDCVAGRQDED 1003

RESULT 44

ABU57552

ID ABU57552 standard; Protein; 1003 AA.

XX

AC ABU57552;

XX

DT 09-APR-2003 (first entry)

DE AIDS associated retrovirus-2 (9B) (ARV-2 (9B)) POL protein.

XX

KW Immunoassay; human immunodeficiency virus; HIV; antibody;

KW gag antigen; acquired immunodeficiency syndrome; AIDS; HIV infection;

KW vaccine; gene therapy; protein expression system; ARV-2;

KW AIDS associated retrovirus 2.

XX

OS Retroviridae.

XX

PN US6458527-B1.

XX

PD 01-OCT-2002.

XX

PF 28-JUN-1993; 93US-0083391.

XX

PR 24-DEC-1987; 87US-0138894.

PR 17-AUG-1992; 92US-0931191.

PR 31-OCT-1984; 84US-0667501.

PR 30-JAN-1985; 85US-0596534.

PR 06-SEP-1985; 85US-0773447.

XX

PA (CHIR) CHIRON CORP.

XX

PI Luciw PA, Dina D, Steimer K, Pescador RS, George-Nascimento C;

PI Parkes D, Halliwell R, Barr PJ, Truett M;

XX

DR WPI: 2003-182063/18.

DR N-PSDB; ABX77297.

XX

PT Immunoassay for detecting the presence of antibodies to a human

PT immunodeficiency virus (HIV) for diagnosing, treating or preventing HIV

PT infection, comprises contacting the sample with an HIV gag antigen -

XX

PS Disclosure; Fig 5B-E; 101pp; English.

XX

CC The invention describes an immunoassay for detecting the presence of

CC antibodies to a human immunodeficiency virus (HIV) in a sample by:

CC contacting the sample with an HIV gag antigen that binds anti-HIV

CC antibodies in an acquired immunodeficiency syndrome (AIDS) patient sera;

CC and determining binding, where the improvement comprises employing as

CC the gag antigen either a synthetic polypeptide or a recombinant

CC polypeptide. The method is used for detecting the presence of antibodies

CC to a human immunodeficiency virus (HIV) in a human sample. The

CC method is used in the preparation of a medicament for diagnosing,

CC treating or preventing HIV infection e.g. a vaccine or for gene

CC therapy. This is the amino acid sequence of a retroviral protein

CC isolated from AIDS associated retrovirus 2 (ARV-2) or HIV.

XX

SX Sequence 1003 AA;

Query Match 57.0%; Score 3109; DB 24; Length 1003;

Best Local Similarity 56.8%; Pred. NO. 8.2e-241;

Matches 565; Conservative 179; Mismatches 243; Indels 12; Gaps 6;

RESULT 44

AAU084391

ID AAU084391 standard; Protein; 995 AA.

XX

Qy 38 GSTGEIYAAREKTERAERETIQ--GSDRGLTAPRAGGDTIQAGTNRGLAAPQFSLWKR 94

Db 11 GKARESSQTRANSFTRRELQVWGGENSL--EAGADR-QGTVSFNF--POITLWQRP 65

Qy 95 VVTAYJEGQVEVLLDTGADDSIVAGIELGNNSYSPKIVGGIGGINTKYEKNVEIEVLNK 154

Db 66 LVTIRIGQLKEALLDTGADDTVLEEMNLPGKWKMGXGXFIVRQVDQIPVEICGH 125

Qy 155 KVRATMTGDTPINIFGRNLTALGMSLNLPAKVEPIKMLPGKDGKPLKROWPLTKK 214

Db 126 KAIGTVLVGPTPVNITGRNLTQIGCTLNFPISPIETVPVKLPKPGMDGPKVKOMPLTEK 185

Qy 215 IEALKEICEKMEKEGEEAPPNTNPTFAIKKDKNKRMLIDIFRELNKVYQDQFTEI 274

Db 186 IKALVEICTEMEKEGKISKIGPENPYNTPVFAIKKKDKTKWRKLVDFRELNRKTQDFEW 245

Qy 275 QLGIPHPAGLAKRRITVLDVGDAYSIPILHEDFRPYTAFTLPVSNNAEPKGYIYKVL 334

Db 246 QLGIPHPAGLKKKSVTVLDVGDAYSFVPLDKDFRKYTAFTIPSIINNETPGIRYQNVLP 305

Qy 335 QGWKGSFALFORTMROVLEPRKANKDVIIIOYMDLILASDRTDLEHVRVVLQKELLN 394

Db 306 QGWKGSFALFQSSMTKILFPFRKQNPDIYVQIMDDLYVGSDELICQHTKTEELRQHLL 365

Qy 395 GLGFSTPDEKFOKPPYHMGYELWPTKWKLOQLPQKEIWTVDIQLKVLGVNMAAQL 454

Db 366 RWGFTTPDKKHOKEPFPLMWGYELHPDKWTVPIMLPEKDSMTVNDIQLKVLGNWASQI 425

Qy 455 YPGIKTKHLRLIRKMTLITEVQWTELAEELEENRIILSQBQEGHYQEKELEATVQ 514

Db 426 YAGIKVKQLCKLLRGTKALTEVPLTEAELEAENREILKEPVHEVYDPSKDLVAEQ 485

Qy 515 KDOENOMTYKIHQEE--KILKVGKYAKVKNTHNTGIRLLAQVYVOKICEALVINGRIPKH 573

Db 486 KOGQGWTYQIYQEPFNKLTGKYMKGHAHNDVKOLTEAVQKVSTESIVIMGKLPKF 545

Qy 574 LPVEREIQWMDNYQWTVIPDWDFVSTPPLVRLAFNLVGDPIPGAETFYTDGSCNRQS 633

Db 546 LPIQKETWEAWMEYQATWIPDWDFVSTPPLVRLAFNLVGDPIPGAETFYTDGSCNRQS 605

Qy 634 KEGKAGYVTRDGRKDKVKKLEQYTNQOAELEAFAMALTDSGPKVNIIVDSQYVNGISASQ 693

Db 606 KLGKAGYVTRDGRKQKVSADTNNQTELOAHLALQDSGLEVINIVTDSQYALGIIOAQ 665

Qy 694 TESEKIVNQITEEMIKKBAIYVAVYPAHKGIGGNOEVDHLVSOGRVLFLEKIPPAQE 753

Db 666 DKSESELVSQILEOLIKKEKYLAVYPAHKGIGGNOEVDHLVSOGRVLFLEKIPPAQE 725

Qy 754 EHEKYHSNVELSHKFGIPNLVARQIVNSCAQOQKEAIGHQVNAELGTWMDCTHLEG 813

Db 726 EHEKYHSNWRMASDFNLPPVAKETIVASCDCQLKGEAMHGQVDCSPGIWQLDCTHLEG 785

Qy 814 KIIVAVHVASGFIEAEVPOESGROTALFLKLASRPVTHLHTDNGANFTSQEVKMA 873

Db 786 KIIVAVHVASGYIEAEVPAETGQETAYFLKLKAGRPVKTTHTDNGSNFTTYYKAA 845

Qy 874 WWTIGESFGVPYNQSGVVEAMNHLKNQISRIREQANTIEIVLMATHCNFKRRRG 933

Db 846 WTAGIKQEFGIPYNQSGVVEAMNHLKNQISRIREQANTIEIVLMATHCNFKRRRG 905

Qy 934 IGDMPSERLINMITTEQETQFLQAKNSKLKDFRVYFREGDRDLWGPGLLWKGE 993

Db 906 XGYSAGERIVDIATDIQTKELQKQIKIQRNFRVYRDRDPVWKGPAKLLWKGE 965

Qy 994 KVGTDIKIIPRKAIRIDYGGQEMDGSGLHLEGAR 1032

Db 966 IQDNDSIKVPRKAKIIRIDYGKQ-----AGDDCVASRQD 1001

infection inhibitor; AIDS associated retrovirus; p24; gp41.

Human lymphotropic virus type III.

Key Location/Qualifiers

Misc-difference 50 /label= Leu, Pro

Misc-difference 615 /label= Lys, Arg

Misc-difference 956 /label= Asp, Asn

US5853978-A.

29-DEC-1998.

29-JUL-1994; 94US-0282857.

08-MAY-1986; 86US-0861016.

24-DEC-1984; 84US-0685272.

04-DEC-1985; 85US-0805069.

02-AUG-1988; 88US-0227568.

19-NOV-1992; 92US-0979391.

29-SEP-1993; 93US-0129009.

29-JUL-1994; 94US-0282857.

(GETH) GENENTECH INC.

Berman PW, Capon DJ, Lasky LA;

WPI; 1999-094894/08.

N-PSDB; AAV74271.

New AIDS-associated fusion poly:peptide(s) - used for the detection of AIDS or to inhibit infection by AIDS associated retrovirus or dissemination of such retrovirus in infected individuals

Example 1; Fig 2A-J; 47pp: English.

This invention describes a novel fusion polypeptide which comprises (a) a first polypeptide sequence of an AIDS associated E', env, or gag polypeptide, that specifically binds complementary antibody and (b) a second polypeptide sequence which is not an AIDS associated virus polypeptide. Also described are (1) a fusion polypeptide having a herpes simplex virus (HSV) gD signal peptide sequence fused in reading frame with a polypeptide sequence other than HSV gD, (2) a nucleic acid encoding a fusion polypeptide as in (1); (3) an expression vector comprising a nucleic acid as in (2); (4) a host comprising a vector as in (3). The AIDS-associated fusion polypeptides can be used in the detection of and vaccination against viral etiological agents of AIDS. They can also be administered as a pharmaceutical agent to inhibit infection by AIDS associated retrovirus or dissemination of such retrovirus in infected individuals.

Sequence 1012 AA;

Query Match 56.9%; Score 3104.5; DB 20; Length 1012;
Best Local Similarity 56.0%; Pred. No. 1.9e-240;
Matches 567; Conservative 179; Mismatches 250; Indels 17; Gaps 5;
34 GSSSGTGEIYAARETERAERTIQGSDRGUTAPRA-----GGD---TIQAFNRLG 83
1 GRSAFLQKAREFSSEQTRANSRRELQVWGRDNNSXSEAGADROGT 60
84 AA---PQFSIMKRPVTVATIEGOPYVELDGTGADDSIVAGIELGNNSPKIVGGIGGF 140
61 VSNFFPQITLWQRPVLTIKIGGOLKEALLDGTGADDTVLEMSLPGKWKPKMIGGIGGF 120
141 TEYKRVVEVLNKKVRATIMTGDPTINIFGRNLTALGMSLNPVAKVEPIKIMLKPGK 200
121 VROYDQILIEICGHAIGVILVGPVTPVNIIGNRLTQIGCTLNFPISTPIETVPVKLPGM 180
201 DGPKLQWPLTKEKEBALKECEKMEKQLEAEAPNTNPTPTFAIKKKKKNKWRMLD 260

181 DGPVKQWPLTEKIKALVEICTEMEKEGKISKIGPENPYNTPVFAIKKKKSTKWRKLV 240
261 FRELNKVTODFTETIOLGIPHPAGLAKRRITVLDVGDAYFSPLHEDFEPYTAFTPSVN 320
241 FRELNKARTQDFWEVQVIGIPHPAGLAKRRKSVTVLDVGDAYFSPLHEDFEPYTAFTPS 300
321 NAEPPGKRYIYKVLPGQWKGSPAIFOHTMRQVLEPFRKANKDVIIIOYMDLIIASDRTDL 380
301 NETPGIRYOYNVLPQWKGSPAIFOSSMTKILEPFRKQNPDIYIYMDLVGSDLEIG 360
381 EHDVVVLQKELLNGLGFTSPDEKFOKDPYHMGYELWPTKWKLOKIQLPQKEIYTVND 440
361 QHRTKTEELRHLRLMGLTTPDKKHQKEPFLMWMGYELHPDKWTQPIVLPEKDSWTVD 420
441 IOKLVGLNWAQLYPGIKTKHLRGLRGKMTLTERVQWTELAEELEENRIILSQEQEG 500
421 IOKLVGLNWAQLYPGIKTKHLRGLRGKMTLTERVQWTELAEELEENRIILSQEQEG 480
501 HYQEEKELEATVQKQDENQWYKIHOEE-KILKVGKYAKVKNTHNGIRLLAAVQYKIG 559
481 VYDPSKDLIAEIQKOGQGWYQIYQEPFKNLKTGKYARMGAHTNDVKQTEAQKIT 540
560 KEALVTWGRIPKPHLPVEREINQWQWNTWQVTPDWDVSVTPPLVRLAFNLVGPDPG 619
541 TESIVIWGKTPKFLPIQKETWETWYQATWPEWEFVNTPPVLKWLWYQLEKEPIVG 600
620 AETFTYDGCNRSQKSGKAGYVTDGDKVKKLEQTTNOQAELEAFAMALTDGSPKVNI 679
601 AETFTYDGCNRSQKSGKAGYVTDGDKVKKLEQTTNOQAELEAFAMALTDGSPKVNI 660
680 VDSQYVMGISASQPTSESKIVNOIIEEMIKKEAIYVAVWPAHKGIGGQVNDHLYSQGI 739
661 TDSQYALGIIQAPDQSESELYNQIIEQLIKKEKYLAWPAHKGIGGQVNDHLYSQGI 720
740 RQVLFLEKTEPAQOEHEKYSNVELSHKFGIPNLVQIVNSCAQCKQGEAIGHGVNA 799
721 RKILFLDGDIDKAQDEHEKYSNWRAMASDFNLPPVVAKEIVASCDCQLKGEAMHGQVDC 780
800 ELGTWQMDCTHLEGKIIIVAVHVASGFIEAEVPIQESGRQTALFLKLLASRPITHLHTD 859
781 SPGIWQDCTHLEGKIIIVAVHVASGFIEAEVPIQESGRQTALFLKLLASRPITHLHTD 840
860 NGANFTSQEVKVMWVWIGIEQSGFVYPNPQSGVVEAMNHLKNQISRIREQANTTETIV 919
841 NGSNFTSATVKAACWAGIKQEFIPYNPQSGVVEAMNHLKNQISRIREQANTTETIV 900
920 LMAIHCNMFRRGGIGDMTPSPERLINMITTEQEIQLQAKNSKLDKDFRVYFRGRDQLWK 979
901 QMAVFTIHNFRRKGGIGYSAGERIVDIIATDIQTKELQKQITKIONFRVYVYRDSRPLWK 960
980 GPELLWKGEGAVLVKVGTDIKTIIPRRKAKIIRDYGGROEMDSGSHLEGARD 1032
961 GPKALLWKGEGAVVIODNSDIKVVPRRKAKIIRDYGKQM---AGDDCVASROD 1010

Search completed: September 25, 2003, 12:02:16

Job time : 64 secs

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	5075.5	93.1	3080	6	52234323-4	Patent No. 52234323	Sequence 5, Appli
2	5057.5	92.8	1055	2	US-08-659-251-5	Sequence 5, Appli	Sequence 5, Appli
3	5057.5	92.8	1055	3	US-09-256-490-5	Sequence 5, Appli	Sequence 5, Appli
4	5057.5	92.8	1055	5	PCR-US96-11445-5	Sequence 5, Appli	Sequence 5, Appli
5	3324	61.0	1018	4	US-09-206-551-46	Sequence 46, Appl	Sequence 46, Appl
6	3301	60.5	1016	4	US-09-625-973-24	Sequence 24, Appl	Sequence 24, Appl
7	3155	57.9	1003	2	US-07-743-357-9	Sequence 9, Appli	Sequence 9, Appli
8	3150.5	57.8	1014	4	US-09-319-588C-6	Sequence 6, Appli	Sequence 6, Appli
9	3119	57.2	1004	2	US-07-743-357-7	Sequence 7, Appli	Sequence 7, Appli
10	3117	57.2	1005	2	US-07-743-357-1	Sequence 1, Appli	Sequence 1, Appli
11	3116	57.2	1016	2	US-07-743-357-3	Sequence 3, Appli	Sequence 3, Appli
12	3112	57.1	1003	4	US-09-309-573-17	Sequence 17, Appl	Sequence 17, Appl
13	3112	57.1	1015	3	US-08-463-210-9	Sequence 9, Appli	Sequence 9, Appli
14	3112	57.1	1015	3	US-09-124-900-3	Sequence 3, Appli	Sequence 3, Appli
15	3112	57.1	1016	2	US-07-743-357-2	Sequence 2, Appli	Sequence 2, Appli
16	3109	57.0	1016	2	US-07-743-357-5	Sequence 5, Appli	Sequence 5, Appli
17	3107	57.0	1016	2	US-07-743-357-4	Sequence 4, Appli	Sequence 4, Appli
18	3105.5	57.0	1003	2	US-07-743-357-8	Sequence 8, Appli	Sequence 8, Appli
19	3095.5	56.8	1003	2	US-07-743-357-10	Sequence 10, Appl	Sequence 10, Appl
20	2999	55.0	913	2	US-07-743-357-22	Sequence 22, Appl	Sequence 22, Appl
21	2957	54.2	913	2	US-07-743-357-6	Sequence 6, Appli	Sequence 6, Appli
22	1921.5	35.2	562	3	US-09-117-217-14	Sequence 14, Appl	Sequence 14, Appl
23	1921.5	35.2	562	4	US-09-735-487-14	Sequence 14, Appl	Sequence 14, Appl
24	1889.5	34.7	560	4	US-09-752-652-1	Sequence 1, Appli	Sequence 1, Appli
25	1810	33.2	1150	3	US-09-238-303-9	Sequence 9, Appli	Sequence 9, Appli
26	1810	33.2	1150	4	US-09-946-239-9	Sequence 9, Appli	Sequence 9, Appli
27	1673	30.7	1031	4	US-08-811-682-15	Sequence 15, Appl	Sequence 15, Appl

```
Qy 481 ELAEAELEENRIILSQBQEGHYQEEKELEATVQKQENQWTKYKHQEKILKVGKYAKV 540
D 1063 ELAEAELEENRIILSQBQEGHYQEEKELEATVQKQDNQWTKYKHQEKILKVGKYAKI 1122
Qy 541 KNTHTNGIRLLAQVQKIGKEALVWGRIPKFLPVERETWEQWMDNYQWTVIPDWFV 600
D 1123 KNTHTNGVRLLAQVQKIGKEALVWGRIPKFLPVERETWEQWMDNYQWTVIPDWFV 1182
Qy 601 STPLVRLAFNLVGDPIPGAETFTDGSNRSQKQEGKAGVYVTDGRKDKVKKLQTTNQQA 660
D 1183 STPLVRLAFNLVGDPIPGAETFTDGSNRSQKQEGKAGVYVTDGRKDKVKRLQTTNQQA 1242
Qy 661 ELEAFAMALTDSPKVNIIVDSQYVMGISQTESKIVNQIIEEMIKKEAIYVAVWP 720
D 1243 ELEAFAMAVTDSGPKVNIVDSQYVMGIVTQPAESESRIYVAVWP 1302
Qy 721 AHKGIGNQEVNHLVSGIRQVLFLEKIEPAQEHEKYSNVKELSHKFGIPNLVARQIV 780
D 1303 AHKGIGNQEVNHLVSGIRQVLFLEKIEPAQEHEKYSNVKELSHKFGIPNLVARQIV 1362
Qy 781 NSCAQCOQKEATHGQVNAELGTWQMDCTHLEKLIIVAVHVASGFIEAEVPOESGRQT 840
D 1363 NTCAQCOQKEATHGQVNAELGTWQMDCTHLEKLIIVAVHVASGFIEAEVPOESGRQT 1422
Qy 841 ALFLLKLASRPITHLHTDNGANFTSOEVKVAWVWIGIEQSGFVYVNPQSQGVVEAMNHH 900
D 1423 ALFLLKLASRPITHLHTDNGANFTSOEVKVAWVWIGIEQSGFVYVNPQSQGVVEAMNHH 1482
Qy 901 LKQISIRIQANTETIYVLMATHCMNFKRRGIGDMTPSERLINMITTEQETQFLOAKN 960
D 1483 LKQIERIRIQANTMETIYVLMATHCMNFKRRGIGDMTPSERLINMITTEQETQFLOAKN 1542
Qy 961 SKLKDPRVYFREGDQLWQPGELLWKGEGAVLVKVGTDIKIIPRRKAKIIRDYGGROEM 1020
D 1543 SKLKNFYFREGDQLWQPGELLWKGEGAVLVKVGTDIKIIPRRKAKIIRDYGGROEM 1602
Qy 1021 DSGSHLEGAREDEGMA 1036
D 1603 DSGSHLEGAREDEGMA 1618
```

RESULT 2

```
US-08-659-251-5
; Sequence 5, Application US/08659251
; Patent No. 5883081
; GENERAL INFORMATION:
; APPLICANT: Kraus, Guenter
; APPLICANT: Wong-Staal, Flossie
; APPLICANT: Talbott, Randy
; APPLICANT: Poeschla, Eric
; TITLE OF INVENTION: Isolation of No. 5883081el HIV-2 Proviruses
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/659,251
; FILING DATE: No. 5883081 yet assigned
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/001,441
; FILING DATE: 26-JUL-1995
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: Garrett-Wackowski, Eugenia
; REGISTRATION NUMBER: 37,330
; REFERENCE/DOCKET NUMBER: 02307E-056410US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1055 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..1055
; OTHER INFORMATION: /note= "pol protein encoded by HIV-2KR"
; US-08-659-251-5
```

```
Query Match 92.8%; Score 5057.5; DB 2; Length 1055;
Best Local Similarity 91.6%; Pred. No. 0;
Matches 949; Conservative 43; Mismatches 43; Indels 1; Gaps 1;

Qy 1 TGRFFRTGPGKCAPOLPRGPPSAGADTNTSPSSSGSGTGEIYAAREKTERAEREFTIQG 60
D 21 TGWFFRDWPMGKEASQLPRDPSPAGADTNTSPSRSPAREVLAAREAEAEENETIQG 80
Qy 61 SDRLGLAPRAGGTIOGATNRGLAAPQFSLWKPVVVYATIEGQPVVLLDTGADDSIVAG 120
D 81 GDRLGLAPRTRTDTORG-DKGFAPQFSLWKPVVVYATIEGQPVVLLDTGADDSIVAG 139
Qy 121 IELGNYSKPIVGGIGGIFINTKEYKNVEIEVLNKKVRATIMTGDTPINIFGRNLTALGM 180
D 140 IELGSNYSKPIVGGIGGIFINTKEYKNVEIKVLNKKVKATIMTGDTPINIFGRNLTALGM 199
Qy 181 SLNLPVAKVBPIMKLPKQDGPGLKQWPLTKKEIALKEICEKMEKEGQLEAPPTNPY 240
D 200 SLNLPVAKVDPIKVLKPGKDGPKVQWPLTKKEIALKEICEKMEKEGQLEAPPTNPY 259
Qy 241 NPTFFAIKKKDKNKRMLIDFRELNVQDFTFQIQLGIPHPAGLAKRRITVLDVGDYF 300
D 260 NPTFFAIKKKDKNKRMLIDFRELNVQDFTFQIQLGIPHPAGLAKRRITVLDVGDYF 319
Qy 301 SIPLHEDFRPYTAFTLPSVNNABPGKRYIKVLPQGMKGSIPAIFQHTMRQVLEPFRKAN 360
D 320 SIPLHEDFRPYTAFTLPSVNNABPGKRYIKVLPQGMKGSIPAIFQHTMRQVLEPFRKAN 379
Qy 361 DVIIIOYDDILLIASDRTDLEHDRVVLQKLKELLNGLGFSPTDDEKFOKDPYHMGYELWP 420
D 380 DVILVOYDDILLIASDRTDLEHDRVVLQKLKELLNGLGFSPTDDEKFOKDPYHMGYELWP 439
Qy 421 TKWKLOKIQLPQKEIWTVDIOKLVLGNWAAOLYPGIKTKHLCLIRIGKMTLTVEQW 480
D 440 TKWKLOKIQLPQKEIWTVDIOKLVLGNWAAOLYPGIKTKHLCLIRIGKMTLTVEQW 499
Qy 481 ELAEAELEENRIILSQBQEGHYQEEKELEATVQKQDNQWTKYKHQEKILKVGKYAKV 540
D 500 ELAEAELEENRIILSQBQEGHYQEEKELEATVQKQDNQWTKYKHQEKILKVGKYAKI 559
Qy 541 KNTHTNGIRLLAQVQKIGKEALVWGRIPKFLPVERETWEQWMDNYQWTVIPDWFV 600
D 560 KNTHTNGVRLLAQVQKIGKEALVWGRIPKFLPVERETWEQWMDNYQWTVIPDWFV 619
Qy 601 STPLVRLAFNLVGDPIPGAETFTDGSNRSQKQEGKAGVYVTDGRKDKVKKLQTTNQQA 660
D 620 STPLVRLAFNLVGDPIPGAETFTDGSNRSQKQEGKAGVYVTDGRKDKVKRLQTTNQQA 679
Qy 661 ELEAFAMALTDSPKVNIIVDSQYVMGISQTESKIVNQIIEEMIKKEAIYVAVWP 720
D 680 ELEAFAMALTDSPKVNIIVDSQYVMGISQTESKIVNQIIEEMIKKEAIYVAVWP 739
Qy 721 AHKGIGNQEVNHLVSGIRQVLFLEKIEPAQEHEKYSNVKELSHKFGIPNLVARQIV 780
D 1055 AHKGIGNQEVNHLVSGIRQVLFLEKIEPAQEHEKYSNVKELSHKFGIPNLVARQIV 1055
```

Db 740 AHKGIGNQVDHLVSQGIQVLFLEKTEPAQEHEEHKYSNVKLSHRFGLPKLVARQIV 799
QY 781 NSCAOQKGEAIGHGVNAELGTWQMDCTHLEGGKIIIVAVHVASFIEAEVIPSQSGROT 840
Db 800 NTCACQKGEAIGHGVNAELGTWQMDCTHLEGGKIIIVAVHVASFIEAEVIPSQSGROT 859
QY 841 ALFLKLASRPITHLHTDNGANFTSQEVKVMVAMWIGIEQSGFVYPNPSQSGVVEAMNHH 900
Db 860 ALFLKLASRPITHLHTDNGANFTSQEVKVMVAMWIGIEQSGFVYPNPSQSGVVEAMNHH 919
QY 901 LKNQISRIREQANTTETIVLMAHOMNFKRRGGIGDMTPSERLINIMITEQBIQLOAKN 960
Db 920 LKNQISRIREQANTTETIVLMAHOMNFKRRGGIGDMTPSERLINIMITEQBIQLOAKN 979
QY 961 SKLKDFRVYFREGDQLWKGPELWKGEGAVLVKVGTDIKIIPRRKAKIIRDYGGROEM 1020
Db 980 SKLKNFRVYFREGDQLWKGPELWKGEGAVLVKVGTDIKIIPRRKAKIIRDYGGREV 1039
QY 1021 DSGSHLEGAREDEGEMA 1036
Db 1040 DSSSHLEGREDEGEVA 1055

RESULT 3

US-09-256-490-5
; Sequence 5, Application US/09256490
; Patent No. 6235881
; GENERAL INFORMATION:
; APPLICANT: Kraus, Guenter
; APPLICANT: Wongs-Staal, Flossie
; APPLICANT: Talbott, Randy
; APPLICANT: Poeschla, Eric
; TITLE OF INVENTION: Isolation of No. 6235881el HIV-2 Proviruses
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/256,490
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/659,251
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Garrett-Wackowski, Eugenia
; REGISTRATION NUMBER: 37,330
; REFERENCE/DOCKET NUMBER: 02307E-056410US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1055 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..1055
; OTHER INFORMATION: /note= "pol protein encoded by HIV-2KR"
US-09-256-490-5

Query Match 92.8%; Score 5057.5; DB 3; Length 1055;
Best Local Similarity 91.6%; Pred. No. 0;
Matches 949; Conservative 43; Mismatches 43; Indels 1; Gaps 1;
QY 1 TGRFFRTGPLGKEAPQLPRGPSSAGADTNSTSGSSSGTGEIYAAREKTERAERTIQQ 60
Db 21 TGWFFRDWPMGKEASQLPDRPSPAGADTNSTSRSPAREVLAAREAEARENETIQG 80
QY 61 SRGLTAPRAGDGTIGATNRCGLAAPQSLWLRPVVTAYIEGQPEVLLDTGADDSIYAG 120
Db 81 GDRGTAPTRRDITQRG--DRGFAAPQFSLWLRPVVTAYIEGQPEVLLDTGADDSIYAG 139
QY 121 IELGNYSPIKIVGGIGGFINTKEYKNVEIEVLNKKVRATIMTGDPIINIFGNILTALCM 180
Db 140 IELGNYSPIKIVGGIGGFINTKEYKNVEIKVLNKKVKATIMTGDPIINIFGNILTALCM 199
QY 181 SUNLPVAKVEPIKIMLPGKDGPKLRQWPLTKETKEALKEICEKMEKEQLEBEPPTNPY 240
Db 200 SUNLPVAKVDPIKIVILKPGKDGPKVRQWPLTKETKEALKEICEKMEREGQLEAPPTNPY 259
QY 241 NPTTFAIKKKDKKWRMLIDFRELANKVTQDFTFEOIGIPHPAGLAKKRITVLDYDAVF 300
Db 260 NPTTFAIKKKDKKWRMLIDFRELANKVTQEFTEIOIGIPHPAGLAKKRITVLDYDAVF 319
QY 301 SIPLHEDFRPYTAFTLPSVNNAPGKRYIKVLPQGWKGSIPAIFQHTMRQVLEPFRKAN 360
Db 320 SIPLHEDFRPYTAFTLPTVNNAPGKRYIKVLPQGWKGSIPAIFQHTMRQVLEPFRKAN 379
QY 361 DVIILQYMDDIILIASDRTDLEHDRVVLQKELLNGLGFSTPDEKFOKDPYPYHMGYELWP 420
Db 380 DVIILQYMDDIILIASDRTDLEHDRVVLQKELLNGLGFSTPDEKFOKDPYPYHMGYELWP 439
QY 421 TKWKLQKIQLPQKEVNTVNDIQKLVGLNWAQAQYPGIKTKHLCLRIEKGKMTLSEVQMT 480
Db 440 TKWKLQKIQLPQKEVNTVNDIQKLVGLNWAQAQYPGIKTKHLCLRIEKGKMTLSEVQMT 499
QY 481 ELAAEAELENRIILSQEQEGHYIYQBEKELEATVQKQDNQNTYKIHOGEKILKVAKYAK 540
Db 500 ELAAEAELENRIILSQEQEGHYIYQBEKELEATVQKQDNQNTYKIHOGEKILKVAKYAK 559
QY 541 KNTHTNGIRLLAQVYQKIGKEALVITWGRIPKPHLPVEREIEWQWMDNTHQVWIPIWDV 600
Db 560 KNTHTNGIRLLAQVYQKIGKEALVITWGRIPKPHLPVEREIEWQWMDNTHQVWIPIWDV 619
QY 601 STPLVLRLAFNLVGDPIPGAETFTYDGSNROSKEGKAGYVTDGCKDKVKLEQTTNQA 660
Db 620 STPLVLRLAFNLVGDPIPGAETFTYDGSNROSKEGKAGYVTDGCKDKVKLEQTTNQA 679
QY 661 ELEAFAMALTDSPKVNIIIVDSQYVYMGISASOPTESKIVNQIIEEMIKKEAIVYAWVP 720
Db 680 ELEAFAMALTDSPKVNIIIVDSQYVYMGISASOPTESKIVNQIIEEMIKKEAIVYAWVP 739
QY 721 AHKGIGNQVDHLVSQGIQVLFLEKTEPAQEHEEHKYSNVKLSHRFGLPKLVARQIV 780
Db 740 AHKGIGNQVDHLVSQGIQVLFLEKTEPAQEHEEHKYSNVKLSHRFGLPKLVARQIV 799
QY 781 NSCAOQKGEAIGHGVNAELGTWQMDCTHLEGGKIIIVAVHVASFIEAEVIPSQSGROT 840
Db 800 NTCACQKGEAIGHGVNAELGTWQMDCTHLEGGKIIIVAVHVASFIEAEVIPSQSGROT 859
QY 841 ALFLKLASRPITHLHTDNGANFTSQEVKVMVAMWIGIEQSGFVYPNPSQSGVVEAMNHH 900
Db 860 ALFLKLASRPITHLHTDNGANFTSQEVKVMVAMWIGIEQSGFVYPNPSQSGVVEAMNHH 919
QY 901 LKNQISRIREQANTTETIVLMAHOMNFKRRGGIGDMTPSERLINIMITEQBIQLOAKN 960
Db 920 LKNQISRIREQANTTETIVLMAHOMNFKRRGGIGDMTPSERLINIMITEQBIQLOAKN 979
QY 961 SKLKDFRVYFREGDQLWKGPELWKGEGAVLVKVGTDIKIIPRRKAKIIRDYGGROEM 1020
Db 980 SKLKNFRVYFREGDQLWKGPELWKGEGAVLVKVGTDIKIIPRRKAKIIRDYGGREV 1039
QY 1021 DSGSHLEGAREDEGEMA 1036


```

; ORGANISM: Simian immunodeficiency virus
;
; FEATURE:
;
; OTHER INFORMATION: Deduced amino acid sequence of SIVrcm Pol protein
US-09-206-551-46

```

Query Match	61.0%	Score 3324;	DB 4;	Length 1018;
Best Local Similarity	59.5%;	Pred. NO. 1.8e-275;		
Matches 627;	Conservative 156;	Mismatches 209;	Indels 62;	Gaps 8;

QY	4	FFRTGLPGK-EAPQL-----PRGPSSAGADTNTSPSSSSSGST	40
Db	1	FFFRETPLGGEAEELSFDFLDSLCSRDGEOLPCRRDAKELSEGRGKTETTE-----	53
QY	41	GEIYAAREKTERAERETIQGSDRCLTAPRAGGDTIQGATNRLGAAPFSLWKRPVVYAI	100
Db	54	-----AGREQER-----GSI GSNLPOFALWKRETTTAQI	85
QY	101	EGOPVEVLLDTGADDSIVAGIELGNYSPIKIVGGIGGFINTKEYKNVEIEVLNKKVRATI	160
Db	86	EGQKVEVLLDTGADDIVIEGIELGNWTPKIIGGIGGYINVKYKNCIEIAGRTAHV	145
QY	161	MTGDTPIINIGRMILFALGMSNLPAVAKVEPIKIMLPKGDGPKLRQWPLTKEKIEALKE	220
Db	146	LVGPTPVNIIGRNVLAKLGATLNFPSIQIETIKVELKSGDGPRVKQWPLSKEXIEALTE	205
QY	221	ICEKMEKEGOLBEAPTPNPNYNTTFAIKKKDKKNKWRMLIDFRELNKVTQDFTETQLGIPH	280
Db	206	ICNAMEKEGKISIGPENPNYNTFICFKKDDSTKRWKLVDVFRELNKTDQFFVOLGIPH	265
QY	281	PAGLAKRRRTITVLVDGDVAYESIDPLHEDFRPYTAFTLPVSNAEPGKRYIKVLPQGWKGS	340
Db	266	PGGLKOCERTITVLIDGDAYESCLLYEPFRKYTAFTIPAVANNQGPVRYQYNVLPQGWKGS	325
QY	341	PAIFQTMKQVLPBPFRRKANKDVIIQYMDILLASDRTDLEHDRVVLQKLKELNGLGFST	400
Db	326	PAIFQASANKILOPFREENPDVITIIQYMDDLFVGSDDRTKLEHDKMKIQLRDHLHFWGFET	385
QY	401	PDEKFOKDPYPHWGVELWPTWKLOKIQLOPEKEIWTNVDIQKLVGYNLWAAQLPGIKT	460
Db	386	PDKKFOKDPYLMWGVELHPKSWTVQBIKPEKEEWTNVDIQKLVGKLNWASQIYSGLRT	445
QY	461	KHLCLIRGKMTLTEEYQWTFELAAELEENRIILSQBOEGHYQDEEKELEATVQKQDENQ	520
Db	446	KELCKLIRGAKALDEKVEMTKEAIEEYENKMILKELHGVYDEKKPLVANIQKLEGGQ	505
QY	521	WYTKIHQEE-KILKGVKYAKVKNTHYNGIKLLAQVOKIGKEALVINGRIPKHPLVVERE	579
Db	506	WSYQIEQESGPKLTGKYAKQKTAHTNEIRMLAGLVQKIAKEATVINGRLPTFRLPTERE	565
QY	580	IBWQWNDYQWVTWIPDWDPSTPPLVRLAFNLVGDPIPCAETFYTDGSCNRSQKEKAG	639
Db	566	VWD-WRSQWQWVTWIPDWEFVSTPPLIRJGYNLVKDPDIPGEEVYVDGAANRNSKIKAG	624
QY	640	YVTRDGKDKVKLEOTTNQOAELEAFAMALTDGSPKVNIIVDSQYVNGIISASQPTESK	699
Db	625	YVTRNGKEVKLEETTNGKAELEAVLLALKDQSPKVNIVTDSQYVYGILEAQPDTSOG	684
QY	700	IVNQIITEMIKKEAIVYAWVPAHKGIGGNOEVDHLVSQGRQVLFLEKIEPAQOEHEKYH	759
Db	685	LVTEILINOMIGKEAVLSWVPAHKGIGGNEEVDKLSKVRQVLFLDGIEKAQOEHEKYH	744
QY	760	SNVKELSHKFGIPNLVARQIVNSCAOQOKEAIGHOVNAELGTQWMDCTHLEKKIIVA	819
Db	745	NNWRALAEQFQIPQIVAKEIVAQCQVKGEAIGHOVDAISPQTQWMDCTHLEKKIIVA	804
QY	820	VHVASGFIAEVIPOBSGROTALFLAKIASRWPITHLHTDNGANFTSOEYKVMYAWWTGIE	879
Db	805	VHVASGYIEAEVIPAETGKETAHELLKLAAWPVPVKRLHTDNGANFTSAAVQAVCWMAQIE	864
QY	880	QSGFVDPYNQSQGVSEAMNHHLLKNQISRIREQANTTETIVLMAIHCNMFKRRGGIDWTP	939
Db	865	HAFGVYNPQSQGVSEMNKQLKIIIEQVREQAQKLETAQOMAVLFINFRKKGIGYSA	924

[illegible]

Qy	999	DIKIPRRKAKIIRDYGGRRQEMDSGSHLEGARED	1032
		: : :	
Db	985	DLKVPPRRKAKIIIEY-GRKDVDSEANLAGRQEE	1017

RESULT 6
US-09-625-972-24
Sequence 24, Application US/09625972
Patent No. 6566513
GENERAL INFORMATION:
APPLICANT: GUERTLER, Lutz Gerhard
APPLICANT: HAUSER, Hans Peter
APPLICANT: DONGMO DELOKO, Yvette Beatrice
APPLICANT: ZEKENG, Leopold
APPLICANT: KAPTUE, Lazare
TITLE OF INVENTION: LENTIVIRUS FROM THE GROUP OF IMMUNODEFICIENCY VIRUSES OF DRILL
(MANDRILLUS LEUCOPHAEUS) AND THEIR USE
FILE REFERENCE: 067595/0106
CURRENT APPLICATION NUMBER: US/09/625,972
CURRENT FILING DATE: 2000-07-29
PRIOR APPLICATION NUMBER: DE 199 36 003.0
PRIOR FILING DATE: 1999-08-03
NUMBER OF SEQ ID NOS: 57
SOFTWARE: PatentIn version 3.0
SEQ ID NO 24
LENGTH: 1016
TYPE: PRT
ORGANISM: SIV - viral
US-09-625-972-24

Query Match	60.5%;	Score 3301;	DB 4;	Length 1016;
Best Local Similarity	59.5%;	Pred. No. 1.7e-273;		
Matches 613;	Conservative 157;	Mismatches 242;	Indels 18;	Gaps 7;

Qy	4	FRTGPIGK-BAPOLPRGPSSAGADNTSPGSSSGSTGCIYIAAREKTERABRETIQSD	62
Db	1	FFRECSJGOWQTBL---SCRATDPNGTTPAIRGSPRE-DATGLHEEGADEGREGSQ	55
Qy	63	RGLTAPRAGGDTIOGATNRLGAAQFSLWKKRPVVYATIEGQPEVLLDTGADDSIVAGIE	122
Db	56	TG-----EGQRLRG---GLQLPQFSLWNRPTTVVEIEGQVEALLDTGADTVLKDLQ	106
Qy	123	LGNNYSKPIVGGIGGINTKYEYKNVEIEVLNKKVRATIMTGDTPPINFGRNILITALGMSL	182
Db	107	LITGNMKFOIITGGIGGATRVKQYFNCKITPVAGKSTHASVLVGTPPNIIIGRNVLKLGCTL	166
Qy	183	NLPVAKYEPRIKMLPGKDGPKLRQWPLTKKIEALKICEKMEKEGOLLEEAPPNPNT	242
Db	167	NFLPSKLTETVKVTLPKGTDPRIQKWPPLSKKEKILALQELCINOMEKEGISRIGPENPNT	226
Qy	243	PTFAIKKKDKNNKRWMLIDFRELNKVTDQTFEIQIGIPHPAGLAKRRITVLVDGDAYFSI	302
Db	227	PVFCIKKKDGASWRKLVDROLNKNVTQDFEVLQIGIPHPGGLKQCEQITVLVDIGDAYFSC	286
Qy	303	PLHEDFRPYTAFTLPSVNNAPBGRYIYKVLPOGWKGSFALFOHTMROVLEPFRKANDV	362
Db	287	PLDEDFRKYTAFTLPSVNNQPGRIYQYNVLPOGWKGSFALFOATADMLKLTFFKEYPEV	346
Qy	363	IIIQYMDILIASDRTLDHRVYVLQKLELNGLGFSTPDEKFKQDPYPYHMMGYELMPTK	422
Db	347	LIIQYMDLFGVSDLNATEHNKMLNKLREHLRFWGLETPDKKKFQKEPPEFWMGVYLHPKK	406
Qy	423	WKLOKIQLPQKEIWTVDIOKLVLGNWAAQLYPCIKTKHLCRLIRGKMTITEEYOWTEL	482
Db	407	WTVOIKQLPEKEQWTVNDIOKLVLGKLNWASQIYSGIKTEJCKLIRGAKPDLDEIWEYTRK	466
Qy	483	AEAELEENRIILSQEIGHYIOEKELEATYVKQDOENQWTKYIHQEE-KILKVGYAKVK	541
Db	467	AELYEENKIIVOEYHGVYOPBKPIMAKVOKLTQOGWYSYOIEOENKPLKAGYARTK	526

Query Match	57.8%;	Score 3150.5;	DB 4;	Length 1014;
Best Local Similarity	56.0%;	Pred. No. 1.3e-260;		
Matches 579;	Conservative 179;	Mismatches 253;	Indels 23;	Gaps 6
Qy	4	FFRTGCLGKAPQALPRGPSSAGADTNTSPSSSGSGTGEIVAAAREKTERAERETIQ-GSD	62	
Db	1	FFR-----EELVSLQRETRKLPPDNKK--ERAHSPATRELWVSGGEHTTGGDAGEPED	53	
Qy	63	RGLTAPRAGGDTTQCATNRRGLAAPQFSLWKRPVYTAIEGQPVVELLDTGADDSIVAGIE	122	
Db	54	RELSVPT-----FNFQITIQWRPVITVTKGEVREALLDTGADTVIELQ	100	
Qy	123	LGNNSPKIVGGIGGFTNKYKWEVTELVNKKVRATIMTGDTPINIFGRNLLTALGMSL	182	
Db	101	LEGKKWKPMIGGIGGFTKVRQYDNITVDIQGRKAVGIVLGGTPVNIIGRNLLTQIGTL	160	
Qy	183	NLPVAKVEPIKIMLPGKDGPKLRQWPLTKKEIKALKEICEMKEGOLEAPTPNPNT	242	
Db	161	NFPISPTETVPVKLPGMDGPKVKQWPLTTKEIALREICTEMEKEGKISRIGPENPYNT	220	
Qy	243	PTFAIKKKDKNKRMLDFRELNKVTDQDFTEIOLGIIHPAGLAKRRRTIVLDVGDAYESI	302	
Db	221	PIFAIKKKDSTKWRKLVDFRELNKRTQDFVEVQLGIIHPAGLQKKKSVTVLDVGDAYESC	280	
Qy	303	PLHEDFRPYTAFTLPSVNNAEPGKRYTYKVLPPQGWKGSPTAFTQTMROVLPPFRKANKDV	362	
Db	281	PLDKDFRKYTAFTIPISNNETPGIRYQYNVLPGQWKGSPTAFTQSTMKILPEFRKHEPEI	340	
Qy	363	IIIQYMDIILIASDRTDLEHDRVVLQKLELLNGLIGFSTPDDEKFOKDPPIYHWMGYELWFTK	422	
Db	341	IIYQYMDLLXVGSDELAQHAQREAVEDRHLLKWGFTTPDKKHQKEPFLWAGYELHPDK	400	

Qy	423	WKLOKIQLOPKEKIWTYVNDIQKLVGLNNAQLOLPGTKIKLCHLRIIRCKMPLTTEBVQNTL 482
Db	401	WTQYIKLPKEDVMTVNDIQKLVGLNNAQLOLPGTKIKLCHLRIIRCKMPLTTEBVQNTL 460
Qy	483	AEAELEENRIILSOEQBGHYQOEKELEATVQKDOENQWYKIHQE-EKILKVGKYAKVK 541
Db	461	AEUELAENREILKEPLUGVYDYGKELVAEIQOGOGOWTYQIYQELUHKNLKTKGYAKMR 520
Qy	542	NTHTGIRLLAQVYQKIGKALYIWRGPRKPHLPVERIEWQWMDNYWQVWTIPDMDVFS 601
Db	521	SAHTNDIKQLVEVVRKYATESIVLWGTGTPKFRPLPVQKEVWEAMWTDHQAOTWIPWEFEVN 580
Qy	602	TPPLVRLAFNLVGDPIGAEFTFYDGCNSROSKEGKAGYVYDRGDKVKKLEOTTNQOAE 661
Db	581	TPPLVKWLYQLETEPIISGAETFFYVDGAANRETKLGKAGFVTDGRQKRVVSIADTITNOKAE 640
Qy	662	LEAFAMALTDGPKVNIIVDSOYVMGISASOPTESSEKIYVNOIIEEMIKKEALYVAVWPA 721
Db	641	LQAILMALQBSGRDVIIVTDSQAMGLIHSQPKSESELVSIIEELIKKERYVLSWPA 700
Qy	722	HKGIGNGOYDHLVSGOIROVLFLFKLEPAQEHEKYSNVKLSHKGFGIPNLVQARQVN 781
Db	701	HKGIGNGOYDHLVSGOIROVLFLFKLEPAQEHEKYSNVKLSHKGFGIPNLVQARQVN 760
Qy	782	SCAQCOQKGBAIIHQVQNAELGTWQMDCTHLEGIIVAVHVASGFIAEAVIPOESGRQTA 841
Db	761	SCDKQLKGEAMHQVNCSPGVQLODCTHLEGIIVAVHVASGFIAEAVIPOESGRQTA 820
Qy	842	LFLILKLASRPITHLTDGNCANFTSQEVKVMVNWIGIEQSGFGVPYNPQSGOVVEAMNHL 901
Db	821	YFLILKLASRPITHLTDGNCANFTSQEVKVMVNWIGIEQSGFGVPYNPQSGOVVEAMNHL 880
Qy	902	KNOISRIREQANTITVLMAIHCMMFKRGGTMDTPSERLINMITTEOEIOFLQAKNS 961
Db	881	KKIIGQIRDOAEHLKTAQVNAVTHNFKRGGIGGYTAGERIIDIATDITQTNLQTOIL 940
Qy	962	KLKDFRVYFREGDQMLKPGCELLWKGEQAVLVKVGTDIKIIPRRKAKIIRDYGGQEMD 1021
Db	941	KVQNFVRYVYRSDRPIWKGPALKLWKGEQAVLVKVGTDIKIIPRRKAKIIRDYGGQEMD 999
Qy	1022	SGSHLEGAREGDEM 1035
Db	1000	DGVASQGDENQEM 1013
RESULT 9		
US-07-743-357-7		
; Sequence 7, Application US/07743357		
; Patent No. 5858646		
; GENERAL INFORMATION:		
; APPLICANT: Kang, Yong C.		
; TITLE OF INVENTION: Polypeptide having immunological		
; TITLE OF INVENTION: activity for use as diagnostic reagent and/or vaccine		
; NUMBER OF SEQUENCES: 22		
; CORRESPONDENCE ADDRESS:		
; ADDRESSEE: KIRBY EADES GALE BAKER		
; STREET: Box 3432, Station D		
; CITY: Ottawa		
; STATE: Ontario		
; COUNTRY: Canada		
; ZIP: K1M 1H8		
; COMPUTER READABLE FORM:		
; MEDIUM TYPE: Floppy disk		
; COMPUTER: IBM PC compatible		
; OPERATING SYSTEM: PC-DOS/MS-DOS		
; SOFTWARE: Patent In Release #1.0, Version #1.30		
; CURRENT APPLICATION DATA:		
; APPLICATION NUMBER: US/07/743,357		
; FILING DATE: 21-AUG-1991		
; CLASSIFICATION: 424		
; PRIOR APPLICATION DATA:		
; APPLICATION NUMBER: PCT/CA90/00062		
; FILING DATE: 23-FEB-1990		


```
ATTORNEY/AGENT INFORMATION:
NAME: Gale, Edwin J.
REGISTRATION NUMBER: 28,584
REFERENCE/DOCKET NUMBER: 30924-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 237-6900
TELEFAX: (613) 237-0045
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1004 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
HYPOTHETICAL: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: Human immunodeficiency virus type 1
STRAIN: SF2
US-07-743-357-7

Query Match 57.2%; Score 3117; DB 2; Length 1004;
Best Local Similarity 56.8%; Pred. No. 6.3e-258;
Matches 567; Conservative 179; Mismatches 241; Indels 12; Gaps 6;

QY 38 GSTGEIYAAREKTERAETRIQ--GSDRLGTAPRAGGDTIQGATNRGLAAPQFSLWKRP 94
Db 11 GKAREFSSEQTRANSPTRELOVWGGENSL--EAGADR-QCTVSFNF--PQITLWQRP 65
QY 95 VVTAYTEGQPVVLLDTGADDSIVAGIELGNYSPIKIVGGIGGIFINTKEYKNVEIEVLNK 154
Db 66 LVIRIGGQLKEALLDTGADDTVLEENMLPGKKRMIGGIGFIRVQYDQIPVETCGH 125
QY 155 KVRATMTGDTPIINIFGRNLTALGMSLNPVAKVEPIKIMLPGKDGPKLRQMLTKK 214
Db 126 KAIGTVLGTPPNIIGRNLTQIGCTLPPIETVPVKLAPGMDGPKVKOMPLTEK 185
QY 215 IEALKEICEKMEKGOLEEAAPNTPTFAIKKDKNKRMLIDFRELNKVYQDFTEI 274
Db 186 IKALVEICTEMEKGKISKIGPENPTYVFAIKKDKSTKRWKLVDRELNKRKTQDFEW 245
QY 275 QLGIPHPAGLAKRRITVDVGDAYSIPLHEDFRTAPTLSVNNAPGKGIYKVL 334
Db 246 QLGIPHPAGLKKKSVTLVDGDAYSIPLHEDFRTAPTLSVNNAPGKGIYKVL 305
QY 335 QGKKGSPAIPOHTMROVLEPRKANKDVIIIOYMDLILASDRDLEHDRVVLQKELN 394
Db 306 QGKKGSPAIPOHTMROVLEPRKANKDVIIIOYMDLILASDRDLEHDRVVLQKELN 365
QY 395 GLGFSTPDERFQDPYHMGYELWPTKWKLOQLPQKEIWTVDIQKLVGLVNWAAQL 454
Db 366 RWGFTTPDKKHQKPEPPFLMWGELHPDKWTVQIPLPEKDSVTVNDIQKLVGLNWSQI 425
QY 455 YPGIKYKHLRLTRGKMTLVEEVQWTELAPAELEENRIILSQBQEGHYOEEKELEATVQ 514
Db 426 YAGIKYKVLKLLRGTKALTEVPLTEAELEAENREILKEPVHEVYDPSKDLVAEIQ 485
QY 515 KQENQWTKYKHOE-KILKVGKAKVKNHTNGIRLLAQVYQKIGKEALVWGRPKPH 573
Db 486 KQCGQNTYQIYEPFNKLTGKYGARMGAHTNDVKOLTEAVQKSTESIVINGKIPKPK 545
QY 574 LPVERIEWQWNYQVWTPDQWDFVSTPPLVRLAFNLVGDPIPGAETFYTDGSCNROS 633
Db 546 LPIQKETWEAWWYQATWIPWEFVNTPLPKLWYQLEKEPIVGAETFYVDGAANRET 605
QY 634 KEGAGYVTRDGKDKVKLEQTTNQAELEAFAMALTDGPKVNIIVDSQYVNGISASQP 693
Db 606 KLGAGYVTRDGRQKSVSADTTNQKTELQAIHLALQDSGLEVNIIVTDSQYALGIQAQ 665
QY 694 TSESKIVNQIIEEMIKKAIYVAVYPAHKIGGNGQVDHLVDSQGTROVLFLEKIPAOE 753
Db 666 DKSESELSQIIEQLIKKVKYLAWPVPAHKIGGNGQVDKLVASAGIRKVLFLNGIDKAOE 725

ATTORNEY/AGENT INFORMATION:
NAME: Gale, Edwin J.
REGISTRATION NUMBER: 28,584
REFERENCE/DOCKET NUMBER: 30924-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 237-6900
TELEFAX: (613) 237-0045
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1005 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
HYPOTHETICAL: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: Human immunodeficiency virus type 1
STRAIN: HXB2
US-07-743-357-1

Query Match 57.2%; Score 3117; DB 2; Length 1005;
Best Local Similarity 56.8%; Pred. No. 9.4e-258;
Matches 564; Conservative 178; Mismatches 248; Indels 6; Gaps 3;
```


QY	38	GSTGEIYAAREKTERAERETIQGSDRGLTAPRAGGDTIQCATNRGLAAPQFSLWKRPPVT	97
Db	12	GRAREESSQTRANSPTRELQWGDNNSPSEAGADRGTVSFNF--FOVTLWQRLPTV	69
QY	98	AVIEGQPVVLLDTGADDSIVAGIELGNNSPKIVGGIGGFINTKEYKNVEIEVLNKKYR	157
Db	70	IKIGQLKEALLDTGADDTVLEEMSLPGRWKPKMIGGIGGFVKRQYDQILIEICGHKAI	129
QY	158	ATIMTGDTPINIFGRNLTALGMSNLNPKVAKPEIKIMLKPGDKGPKLQWPLTRKIEA	217
Db	130	GTVLVGPPTVNIIGNLLTQIGCTLNFPIISPETVPVKLPGMDGPKVQWPLTEKIRA	189
QY	218	LKEICEKMEKEGQLEEAPPNTNPTFAIKKDKKNKRWMLDFRELNNKVTQDFTEIQLG	277
Db	190	LVEICTEMEKEGKISKIGPENYPNTPVFAIKKDKSTKRWKLVDFRELNNKRTQDFEVLG	249
QY	278	IPHPAGLAKKRITVLDVGDAFYSIPLHEDFRPYTAFTLPSVNNAPGKRYIKVLPQGW	337
Db	250	IPHPAGLKKKSVTLVDVGDAFYSVPLDEDFRKYTAFTTIPSINNETPGIRYQYNVLPQGW	309
QY	338	KGSPALFQHTMROVLEPPFRKANKDVIIIOYMDLIIASDRTDLEHDRVVLQKELLNGG	397
Db	310	KGSPALFQSSWTKILEPPFRKQPDVVIYOYMDLVGSDLEIGQHRTKTEELRQHLRWG	369
QY	398	FSTPDEKFOKPPYHMGYELWPTKWKLOIQLPQKEIWTVNDIQKLVGLNWAQLYPG	457
Db	370	LATPDKKHKEPPLWMGYELHPDKWTQPIVLPEKDSWTVNDIQKLVGLNWAQSIYPG	429
QY	458	IKTKHLCLRLIRKMWLTETEVQWTELAELEENRILISOEGHYHQEKELEANTVKDQ	517
Db	430	IKVRQLCKLLRTKALTEVIPITPEAELELAENREILKEPVHGVHVDYPSKDLIAEQKG	489
QY	518	ENQWYKIHQEE-KILKVGKAKVKNTHNGIRLLAQQVVKIGKICALVINGRIPFHLPV	576
Db	490	QGWYQIYQEFKNLKTKYARMGAHTNDVKQTEAVQKTTTESIVINGKTPFKLPI	549
QY	577	EREINQWMDNWQVWTVPDMDFVSTPLRLAFNLGDPPIGAETFTYDGSNQRSKBG	636
Db	550	QKETWETWTEYQATWIPWEFVNTPLVLKWLWYLEKEPIVGAETFFYVDGAANRETKLG	609
QY	637	KAGYVTDGKDKVKKLEQTTNOQAELEAFAMALTDGSGPKVNIIVDSQYVMGISASOPTES	696
Db	610	KAGYVYNRGRQKVVTLTDTTNOKTELOAIYLALQDSGLEVNITVDSQYALGIIQAQPDGS	669
QY	697	ESKIVNQITEEMIKKEAVIVAWPAHKGIGGNOEVDHLVSQIGROVLPLEKTEPAOEHE	756
Db	670	ESELVNIQIELIKEKYLAWPAHKGIGGNEQDKLVSAIGIRKVLFDGIDKQAQDEHE	729
QY	757	KYHSNVKELSHKFGIPNLVARQVNSCAQCOQKGEAIIHQVNAELGTWQMDCTHLEGKII	816
Db	730	KYHSNWRAMASDFNLPPVVAKEIVASCDKQLKGEAMHQVQDCSPGIMQLDCTHLEGKI	789
QY	817	IYAVHVASGFIEAEVIPAQSGRQTALFLLLKASRWPITHLTDNGANFTSQEVKVAWVI	876
Db	790	LVAHVASGYIEAEVIPAETQETAYFLLLKLAGRPVKTITHDNGSNETGATVRAACWMA	849
QY	877	GIEQSGFVPYNPQSGVVEAMHNLKNQISIREQANTTETIVLMAIHCNFKRGGIGD	936
Db	850	GIKQEGIPYNPQSGVVESNMKELKIIIGQVRDQAEHLKTAQVQNAVFHNFKRGGIGG	909
QY	937	MTPSERLINMITTEIQFLOAKNSKLDKFRVYFREGRDQLWKGPGLKGEQAVLVKV	996
Db	910	YSAGERIVDIIATDITQKLEQKITQINFRVYVYRNSRLSWKPAKLLWKGEQAVVIQD	969
QY	997	GTDIKILPRKAKIIRDYGGREMDSGSHLEGARED 1032	
Db	970	NSDIKWVRRKAKIIRDYGKQM---AGDDCVASQD 1002	
RESULT 11			
US-07-743-357-3			
; Sequence 3, Application US/07743357			

;	Patent No. 5858646		
;	GENERAL INFORMATION:		
;	APPLICANT: Kang, Yong C.		
;	TITLE OF INVENTION: Polypeptide having immunological		
;	activity for use as diagnostic reagent and/or vaccine		
;	NUMBER OF SEQUENCES: 22		
;	CORRESPONDENCE ADDRESS:		
;	ADDRESSEE: KIRBY EADES GALE BAKER		
;	STREET: Box 3432, Station D		
;	CITY: Ottawa		
;	STATE: Ontario		
;	COUNTRY: Canada		
;	ZIP: K1M 1H8		
;	COMPUTER READABLE FORM:		
;	COMPUTER: IBM PC compatible		
;	OPERATING SYSTEM: PC-DOS/MS-DOS		
;	SOFTWARE: PatentIn Release #1.0, Version #1.30		
;	CURRENT APPLICATION DATA:		
;	APPLICATION NUMBER: US/07/743,357		
;	FILING DATE: 21-AUG-1991		
;	CLASSIFICATION: 424		
;	PRIOR APPLICATION DATA:		
;	APPLICATION NUMBER: PCT/CA90/00062		
;	FILING DATE: 23-FEB-1990		
;	ATTORNEY/AGENT INFORMATION:		
;	NAME: Gale, Edwin J.		
;	REGISTRATION NUMBER: 28,584		
;	REFERENCE/DOCKET NUMBER: 30924-2		
;	TELECOMMUNICATION INFORMATION:		
;	TELEPHONE: (613) 237-6900		
;	TELEFAX: (613) 237-0045		
;	INFORMATION FOR SEQ ID NO: 3:		
;	SEQUENCE CHARACTERISTICS:		
;	LENGTH: 1016 amino acids		
;	TYPE: amino acid		
;	STRANDEDNESS: not relevant		
;	TOPOLOGY: not relevant		
;	MOLECULE TYPE: protein		
;	HYPOTHETICAL: NO		
;	FRAGMENT TYPE: Internal		
;	ORIGINAL SOURCE:		
;	ORGANISM: Human immunodeficiency virus type 1		
;	STRAIN: BH5		
;	US-07-743-357-3		
;	Query Match	57.2%;	Score 3116; DB 2; Length 1016;
;	Best Local Similarity	56.3%;	Pred. No. 1.2e-257;
;	Matches	564;	Conservative 182; Mismatches 249; Indels 6; Gaps 3;
QY	33	SGSSSGSTGEIYAAREKTERAERETIQGSDRGLTAPRAGGDTIQCATNRGLAAPQFSLWK	92
Db	18	SQTRANSPTISSEQTRANSPTRELQWGDNNSPSEAGADRGTVSFNF--PQITLWQ	75
QY	93	RPVVTAYIEGQPVVLLDTGADDSIVAGIELGNNSPKIVGGIGGFINTKEYKNVEIEVL	152
Db	76	RPLVTKIGQQLKEALLDTGADDTVLEEMSLPGRWKPKMIGGIGGFVKRQYDQILIEIC	135
QY	153	NKKVRATMTGPTPINIFGRNLTALGMSNLNPKVAKPEIKIMLKPGDKGPKLQWPLTK	212
Db	136	GHKALGTVLGPTVNIIGNLLTQIGCTLNFPIISPETVPVKLPGMDGPKVQWPLTE	195
QY	213	EKIEALKEICEKMEKEGQLEEAPPNTNPTFAIKKDKKNKRWMLDFRELNNKVTQDFT	272
Db	196	EKIKALVEICTEMEKEGKISKIGPENYPNTPVFAIKKDKSTKRWKLVDFRELNNRTQDFW	255
QY	273	ETQLGIPHPAGLAKKRITVLDVGDAFYSIPLHEDFRPYTAFTLPSVNNAPGKRYIKV	332
Db	256	EVLQGIPHPAGLKKKSVTLVDVGDAFYSVPLDEDFRKYTAFTTIPSINNETPGSYQYNV	315
QY	333	LPQWKGSPALFQHTMROVLEPPFRKANKDVIIIOYMDLIIASDRTDLEHDRVVLQKEL	392
Db	316	LPQWKGSPALFQSSWTKILEPPFRKQPDVVIYOYMDLVGSDLEIGQHRTKIEELRQH	375

APPLICANT: CHANG, Nancy T.
APPLICANT: GALLO, Robert C.
APPLICANT: WONG-STAL, Flossie
TITLE OF INVENTION: CLONING AND EXPRESSION OF HTLV-III DNA
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463.210
FILING DATE: 05-JUN-1995
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 06/693,866
FILING DATE: 23-JAN-1985
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 06/659,339
FILING DATE: 10-OCT-1984
ATTORNEY/AGENT INFORMATION:
NAME: Serunian, Leslie A.
REGISTRATION NUMBER: 35,353
REFERENCE/DOCKET NUMBER: 2026-4193US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1015 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: HTLV-III
FEATURE:
NAME/KEY: Protein
LOCATION: 1..1015
OTHER INFORMATION: /note= "pol protein of HTLV-III"
US-08-463-210-9

Query Match 57.1%; Score 3112; DB 3; Length 1015;
Best Local Similarity 56.2%; Pred No. 2.6e-257;
Matches 563; Conservative 183; Mismatches 249; Indels 6; Gaps 3;
QY 33 SGSSGSGTGEIYAAREKTERAERTIQGSDRGLTAPRAGGDTIQATNRGLAAPQFSLWK 92
DB 18 SEQTRANSPITSSQETRANSPTRRELQVGRDNNSPSEAGADRGQTVSFNF--PQITLWQ 75
QY 93 RPVTAYTEGQPVVELLDTGADDSIVAGIELGNYSKPIVGGIGGIFINTKEYKNVEIEVL 152
DB 76 RPLVTIKIGQKKEALLTGADDTVLEMSLPGWRKPKMIGGIGGIFKVRQYDQILIEIC 135
QY 153 NKKVRATMTGDTPTINIFGRNLTALGMSLNPVAKVBPITMLKPGKDGPKLRQMLTK 212
DB 136 GHKAGTVLVGTPVPIVIGRNLTLTGCTLNFPIPTVPVKLAPGMDGPKVRQMLPTE 195
QY 213 EKIEALKICEKMEKEGEEAPPTNPTTFAIKKKDKNKRMLIDFRELNKVTDQFT 272
DB 196 EKIKALVEICTEMEKEGKISKIGPENPTYVFAIKKKDKNKRMLIDFRELNKVTDQFW 255
QY 273 EIQLGIPHAGLAKRRITVLVDGDAYFISPLHEDFRPYTAFTLPSVNNAPGKRYIKV 332
DB 256 EVQLGIPHAGLKKKSVTVLVDGDAYFVPLDEDFRKYTAFTIPSINNETPGIRYQYNV 315

QY 333 LPQGWKGSFAIHOHTMROVLEPERKANKDVIIIOYMDDILIASDRTDLEHDRVVLQKL 392
DB 316 LPQGWKGSFAIHOHTMROVLEPERKANKDVIIIOYMDDILIASDRTDLEHDRVVLQKL 375
QY 393 LNLGLFSTPDEKQFQDPYHMGYELWPTKWKLOKIQLPQKEIWTVDNDIKLVGLVNWAA 452
DB 376 LLRWGLTTPDKKHQKPEPPFLWNGYELHDPKWTVPQIVLPKESDWTVDNDIKLVGLNWA 435
QY 453 QLYPGIKTKHLCLIRGKMTLTFEVSQWTELAELAEENRIILISQEQEGHYQYEELEAT 512
DB 436 QIYPGKVRQCKLLRGTKALTEVPLTEAELEAENREILKEPVHGYVYDPSDLIAE 495
QY 513 VOKDOENQWYKIHQEE-KILVGVYAKVKNTHNGIRLLAOWGVKIGKEALVIGRIPK 571
DB 496 IQKQGGQWYQIYQEPFKNLTKGYARMGAHTNDVKQLTAVOKITTESIVIGKTPK 555
QY 572 PHLVREIWEQWMDNYQWTVIPDWFVSTPPLVRLAFNLVGDPIPGAETFTDGSNCR 631
DB 556 FKLPIQKETWETWETWETWETWETWETWETWETWETWETWETWETWETWETWETWET 615
QY 632 QSKECKAGYVTRDGRKDKVKKLEQTTNQQAELAFAMALTDSPKVNIIVDSQYVNGIS 691
DB 616 ETKLGAGYVTRDGRKDKVKKLEQTTNQQAELAFAMALTDSPKVNIIVDSQYVNGIS 675
QY 692 QPTESEKIVNOIIEEMIKKEAIYVAVWPAHKGIGGNOEVDHLVSGIGROVLFLKIEPA 751
DB 676 QPDKSESELVNOIIEOLIKKEKYLAWPAHKGIGGNOEVDHLVSGIGROVLFLKIEPA 735
QY 752 QEEHEKYSNVELSHKFGIPNLVARQIVNSCAQCOQKEATHGOVNAELGTWQMDCTHL 811
DB 736 QDEHEKYSNMRAMASDENLPPVAKIIVASCDKQLKGEAMHGQVDCSPGIWQDCTHL 795
QY 812 ECKIIIVAVHVASGEIEAFVPOESGROTALELLKLSRWPITHLTNGANFTSOEYKM 871
DB 796 ECKIIIVAVHVASGEIEAFVPOESGROTALELLKLSRWPITHLTNGANFTSOEYKM 855
QY 872 VAWWIGIEQSGVYPNPOSGVVEAMNHLKNOISIRERQANTIIETIVLMAHCHNFKRR 931
DB 856 ACWAGIKQEGFIPYPOSGVVEAMNHLKNOISIRERQANTIIETIVLMAHCHNFKRR 915
QY 932 GGIGMDTPSERLINMITTEQETQFLOAKNSKLUKDFRVYFREGRODLKMGPGELLWKGEA 991
DB 916 GGIGYSAGERIVDIATDIQTKELQKQITKIQNFVRYRDSRNLKMGPAKLLWKGEA 975
QY 992 VLVKVTGDIKIIPRKAKIIRDYGGROEMDSHLEGARE 1032
DB 976 VVIQDNDIKVYPRRKAKIIRDYGKQ---AGDDCVASQD 1013

RESULT 14
US-09-124-900-3
Sequence 3, Application US/09124900
Patent No. 6268484
GENERAL INFORMATION:
APPLICANT: KATINGER, Hermann
APPLICANT: BUCHACHER, Andrea
APPLICANT: ERNST, Wolfgang
APPLICANT: BALLAUN, Claudia
APPLICANT: PURTSCHER, Martin
APPLICANT: TRKOLA, Alexandra
APPLICANT: PREDL, Renate
APPLICANT: SCHMATZ, Christine
APPLICANT: KLINA, Annelies
APPLICANT: STEINDL, Franz
APPLICANT: MUSTER, Thomas
TITLE OF INVENTION: HIV-Vaccines
FILE REFERENCE: 1939-112p
CURRENT APPLICATION NUMBER: US/09/124,900
CURRENT FILING DATE: 1998-07-30
PRIOR APPLICATION NUMBER: PCT/EP95/01481
PRIOR FILING DATE: 1995-04-19
NUMBER OF SEQ ID NOS: 11

; SOFTWARE: Patentin version 3.0									
; SEQ ID NO 3									
; LENGTH: 1015									
; TYPE: PR1									
; ORGANISM: Human immunodeficiency virus type 1									
US-09-124-900-3									
Query Match 57.1%; Score 3112; DB 3; Length 1015;									
Best Local Similarity 56.2%; Pred. No. 2.6e-257;									
Matches 563; Conservative 183; Mismatches 249; Indels 6; Gaps 3;									
Qy	33	SGSSGSGTGEIYAAREKTERAERETIQGSDRGLTAPRAGDGTIQGATNRGLAAPQFSLWK	92						
Db	18	SEQTRANSPITSSQTRANSPTRELQVWGRDNNSPEAGADRGQTVSENF--PQITLWQ	75						
Qy	93	RPVVTAYIEGQPVVEVLDTGADDSIVAGIELGNNSPKIVGGIGGFTNKEYKNVEVL	152						
Db	76	RPLVTIKIGQQLKEALLDTGADDTVLEMSLPGRWKPKMIGGIGGFTKVRQYDQILIEIC	135						
Qy	153	NKVRATIMTGDTPIINIFGRNLTALGMSLNPVAKVEPIKMLKPGKDGPKLRQWPLTK	212						
Db	136	GKALGIVLVGPTFVNIIGNRLTQIGCTLNFPISPIETVPVFKLPGMDGPKVQWPLTE	195						
Qy	213	EKIEALKECKMEKEQLEAEAPTNPYPTFAIKKKDKNKRMLIDFRELNKVTDFT	272						
Db	196	EKIALVEICTEMEKEGKISKIGPENPYTPVPAIKKKDKTKWRKLVDFRELNRKTDQF	255						
Qy	273	EIQIGIHPAGLAKKRITVLVDGDAYFSIPLHEDFRPYTAFTLPVSNNAEPEKRIYKV	332						
Db	256	EVQIGIHPAGLKKKSVTVLVDGDAYFSVPLDEDFRKYTAFTIPSNNETPGIRQYVY	315						
Qy	333	LPGWKGSPIAFQHTMRQVLEPRKANKOVIIQYMDLILIASDRTDLEHDRVVLQKEL	392						
Db	316	LPGWKGSPIAFQHTMRQVLEPRKANKOVIIQYMDLILIASDRTDLEHDRVVLQKEL	375						
Qy	393	LNLGFTSPDEKQKOPPYHWMGYELMPTKWLQKQOLPOKEITWVNDIOKLVGLNAA	452						
Db	376	LLRWGLTPDKHOKPEPFPLWNGYELHPDKWTQPIVLPKSDSWTVNDIOKLVGLNAA	435						
Qy	453	QLYPGIKTKHLRLIKMTLTEVOWTELAELAENRILISQEQEGHYVQEEKELEAT	512						
Db	436	QIYPGIKVQLKRLTKALTEVPIETEAELAENRILISQEQEGHYVQEEKELEAT	495						
Qy	513	VQDQENQWYTKHQEE-KILYGVKAKVKNTHNGIRLLAQVQVQKIGKALYVIGRIPK	571						
Db	496	IQKOGQGWYQIYQEPFKNLTKGYARMRGAHTNDVKQLTEAVQKTTESIVIGKTPK	555						
Qy	572	FHLVPERIEQWNDWYQVTPDWDVSTPLVLRLAFNLVGDPIPGAETFTYDGSNR	631						
Db	556	FKLPQIKETWETWTEYQWATWPEWFEVNTPLVLWYQLEKEPIVGAETFYVDGAANR	615						
Qy	632	QSKEGKAGYVTDGKDKVKLEQTTNQAELAFAFAMALTDSPKPVNIIVDSQVVMGISAS	691						
Db	616	ETKLGRAGYVTKGRQVWPLTNTWQTKELQAIYALQDSGLEVNIVTDSQALGIIQA	675						
Qy	692	OPTSESKIYNQIIEEMIKKEATYVAVWPAHKGIGGQNEVDHLVSQIGQVLFLEKIEPA	751						
Db	676	QPKSESELVNIQIEQLIKKEKYLAWVPAHKGIGGQNEVDHLVSQIGQVLFLEKIEPA	735						
Qy	752	QEBHEKHSNVELSHKFGIPNLVQVYNVSCAQCKQGEAIGHQVNAELGTWQMDCTHL	811						
Db	736	QDEHEKHSNWRAMASDFNIPPVYAKEIVASCDQLKGEAMHGQVDCSPGIWQLDCTHL	795						
Qy	812	EGRIIVAVHVASGFTAEVIEPQESGROATFLFKLASRWPITHLTDNGANTTSQEVKM	871						
Db	796	EGKVLVAVHVASGYEAEVIPAETQETAYFLKLKLAGRPVKTITHDNGNSFATVKA	855						
Qy	872	VANWIGEQSGFYPPYQSGVYVAMNHLKNOISRIEQAQNTIETIVLMAIHCMPKRR	931						
Db	856	ACWAGIKQEGFIPYQSGVYVAMNHLKNOISRIEQAQNTIETIVLMAIHCMPKRR	915						
Qy	932	GGIGDMTPSRLLINMTTEQIEQFLQAKNSKLDKFRVYFREGRDQLWKGPGELLWKGEA	991						
; US-07-743-357-2									
Query Match 57.1%; Score 3112; DB 2; Length 1016;									
Best Local Similarity 56.2%; Pred. No. 2.6e-257;									
Matches 563; Conservative 183; Mismatches 249; Indels 6; Gaps 3;									
Qy	33	SGSSGSGTGEIYAAREKTERAERETIQGSDRGLTAPRAGDGTIQGATNRGLAAPQFSLWK	92						
Db	18	SEQTRANSPITSSQTRANSPTRELQVWGRDNNSPEAGADRGQTVSENF--PQITLWQ	75						
Qy	93	RPVVTAYIEGQPVVEVLDTGADDSIVAGIELGNNSPKIVGGIGGFTNKEYKNVEVL	152						
Db	76	RPLVTIKIGQQLKEALLDTGADDTVLEMSLPGRWKPKMIGGIGGFTKVRQYDQILIEIC	135						
Qy	153	NKVRATIMTGDTPIINIFGRNLTALGMSLNPVAKVEPIKMLKPGKDGPKLRQWPLTK	212						
Db	136	GKALGIVLVGPTFVNIIGNRLTQIGCTLNFPISPIETVPVFKLPGMDGPKVQWPLTE	195						
Qy	213	EKIEALKECKMEKEQLEAEAPTNPYPTFAIKKKDKNKRMLIDFRELNKVTDFT	272						
; US-07-743-357-2									
Query Match 57.1%; Score 3112; DB 2; Length 1016;									
Best Local Similarity 56.2%; Pred. No. 2.6e-257;									
Matches 563; Conservative 183; Mismatches 249; Indels 6; Gaps 3;									
Qy	33	SGSSGSGTGEIYAAREKTERAERETIQGSDRGLTAPRAGDGTIQGATNRGLAAPQFSLWK	92						
Db	18	SEQTRANSPITSSQTRANSPTRELQVWGRDNNSPEAGADRGQTVSENF--PQITLWQ	75						
Qy	93	RPVVTAYIEGQPVVEVLDTGADDSIVAGIELGNNSPKIVGGIGGFTNKEYKNVEVL	152						
Db	76	RPLVTIKIGQQLKEALLDTGADDTVLEMSLPGRWKPKMIGGIGGFTKVRQYDQILIEIC	135						
Qy	153	NKVRATIMTGDTPIINIFGRNLTALGMSLNPVAKVEPIKMLKPGKDGPKLRQWPLTK	212						
Db	136	GKALGIVLVGPTFVNIIGNRLTQIGCTLNFPISPIETVPVFKLPGMDGPKVQWPLTE	195						
Qy	213	EKIEALKECKMEKEQLEAEAPTNPYPTFAIKKKDKNKRMLIDFRELNKVTDFT	272						
; US-07-743-357-2									
Query Match 57.1%; Score 3112; DB 2; Length 1016;									
Best Local Similarity 56.2%; Pred. No. 2.6e-257;									
Matches 563; Conservative 183; Mismatches 249; Indels 6; Gaps 3;									
Qy	33	SGSSGSGTGEIYAAREKTERAERETIQGSDRGLTAPRAGDGTIQGATNRGLAAPQFSLWK	92						
Db	18	SEQTRANSPITSSQTRANSPTRELQVWGRDNNSPEAGADRGQTVSENF--PQITLWQ	75						
Qy	93	RPVVTAYIEGQPVVEVLDTGADDSIVAGIELGNNSPKIVGGIGGFTNKEYKNVEVL	152						
Db	76	RPLVTIKIGQQLKEALLDTGADDTVLEMSLPGRWKPKMIGGIGGFTKVRQYDQILIEIC	135						
Qy	153	NKVRATIMTGDTPIINIFGRNLTALGMSLNPVAKVEPIKMLKPGKDGPKLRQWPLTK	212						
Db	136	GKALGIVLVGPTFVNIIGNRLTQIGCTLNFPISPIETVPVFKLPGMDGPKVQWPLTE	195						
Qy	213	EKIEALKECKMEKEQLEAEAPTNPYPTFAIKKKDKNKRMLIDFRELNKVTDFT	272						
; US-07-743-357-2									
Query Match 57.1%; Score 3112; DB 2; Length 1016;									
Best Local Similarity 56.2%; Pred. No. 2.6e-257;									
Matches 563; Conservative 183; Mismatches 249; Indels 6; Gaps 3;									
Qy	33	SGSSGSGTGEIYAAREKTERAERETIQGSDRGLTAPRAGDGTIQGATNRGLAAPQFSLWK	92						
Db	18	SEQTRANSPITSSQTRANSPTRELQVWGRDNNSPEAGADRGQTVSENF--PQITLWQ	75						
Qy	93	RPVVTAYIEGQPVVEVLDTGADDSIVAGIELGNNSPKIVGGIGGFTNKEYKNVEVL	152						
Db	76	RPLVTIKIGQQLKEALLDTGADDTVLEMSLPGRWKPKMIGGIGGFTKVRQYDQILIEIC	135						
Qy	153	NKVRATIMTGDTPIINIFGRNLTALGMSLNPVAKVEPIKMLKPGKDGPKLRQWPLTK	212						
Db	136	GKALGIVLVGPTFVNIIGNRLTQIGCTLNFPISPIETVPVFKLPGMDGPKVQWPLTE	195						
Qy	213	EKIEALKECKMEKEQLEAEAPTNPYPTFAIKKKDKNKRMLIDFRELNKVTDFT	272						

```
Db 196 ERIKALVEICTEMEKEGKISKIPENYNTPTFAIKKDDSTKRWKLVDPRFELNKRQDFW 255
QY 273 EQLGIPHPAGLAKRRITVLDVGDAYSIPLHEDFRPYTAFTLPSVNNAPGKRYIKV 332
Db 256 EYQLGIPHPAGLAKRRITVLDVGDAYSIPLHEDFRPYTAFTLPSVNNAPGKRYIKV 315
QY 333 LPQGWKSPAIQFHTMRQVLEPRFRKANKDVIIIOYMDLIIASDRDLEHDRVVLQKLB 392
Db 316 LPQGWKSPAIQFHTMRQVLEPRFRKANKDVIIIOYMDLIIASDRDLEHDRVVLQKLB 375
QY 393 LNLGFSPTDEKQDPDPYHMGYELWPTKWKQLQIPLPQKEIWTVDIQLVGLVNWAA 452
Db 376 LLRWGLTTPDKKHQKEPPFLWGYELHDPKWTVPVILPEKDSWTVDIQLVGLVNWAA 435
QY 453 QLYPGIKYKHLRGLRGKMTLFEVQWTELAELAEENRIILSQOEGHYOEEKELEAT 512
Db 436 QIYPGIKYKHLRGLRGKMTLFEVQWTELAELAEENRIILSQOEGHYOEEKELEAT 495
QY 513 VOKQOENQWYKIHQEE-KILKVGKAKYKNTHTNGIRLLAQVQKIGKEALVWGRIPK 571
Db 496 IQKQOGQWYQIYQEPKNTKGTGYARMGAHTNDVKQTEAVQKITTESIVGKTPK 555
QY 572 FHLPEREWEQWQWNYQWTVIPDWDVSTPPLVRLAFNLVGDPIPGAETTYTGSCNR 631
Db 556 FKLPIQKETWETWYQWTVIPDWDVSTPPLVRLAFNLVGDPIPGAETTYTGSCNR 615
QY 632 OSKEGKAGVYDRGKVKKLBQTTNQOAELEAFAMALTDGPKYNIIVDSQYVNGISAS 691
Db 616 ETKLKGAGVYDRGKVKKLBQTTNQOAELEAFAMALTDGPKYNIIVDSQYVNGISAS 675
QY 692 QPTESEKIVNQIIEEMIKKEAIYVAVYPAHKGIGGNOEVDHLVSGIROVLFLEKIPBA 751
Db 735 QPDKSESELVQNIIEQLKKEKYLAWYPAHKGIGGNOEVDHLVSGIROVLFLEKIPBA 705
QY 752 QEEHEKYSNVKELSHKGIPLNVARQIVNSCAQOQKGEATHGQVNAELGTWQMDCTHL 811
Db 736 QDEHEKYSNVKELSHKGIPLNVARQIVNSCAQOQKGEATHGQVNAELGTWQMDCTHL 795
QY 812 EKKIIVAVHVASGIEAEVIEPOESGROTALFLKLASRWPTIHLTONGANFTQEVKM 871
Db 796 EKKIIVAVHVASGIEAEVIEPOESGROTALFLKLASRWPTIHLTONGANFTQEVKM 855
QY 872 VAWMIGIEQSGFYVNPQSGVVEAMNHLKNOISRIREQANTIEIVLMAIHCNFKRR 931
Db 856 ACWAGIQEFGIPIPNQSGVVEAMNHLKNOISRIREQANTIEIVLMAIHCNFKRR 915
QY 932 GGIGDMTYSERLINMITTEQETQFLOAKNSKLKDPVYFREGRODLWGPGBELLWKGEA 991
Db 916 GGIGYSAGERIVDIATDIOTKELQKQITKTONFRVYRDSRNLWKPGBELLWKGEA 975
QY 992 VLAVKGTDIKIIPRKAIIIRDYGGQEMDSGSHLEGAREQ 1032
Db 976 VVIQNSDLKVVPRKAKIIRDYGGQEMDSGSHLEGAREQ 1013
```

RESULT 16

```
US-07-743-357-5
; Sequence 5, Application US/07743357
; Patent No. 5858646
; GENERAL INFORMATION:
; APPLICANT: Kang, Yong C.
; TITLE OF INVENTION: Polypeptide having immunological
; TITLE OF INVENTION: activity for use as diagnostic reagent and/or vaccine
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KIRBY EADES GALE BAKER
; STREET: Box 3432, Station D
; CITY: Ottawa
; STATE: Ontario
; COUNTRY: Canada
; ZIP: K1M 1H8
; COMPUTER READABLE FORM:
```

```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/743,357
; FILING DATE: 21-AUG-1991
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/CA90/00062
; FILING DATE: 23-FEB-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Gale, Edwin J.
; REGISTRATION NUMBER: 28,584
; REFERENCE/DOCKET NUMBER: 30924-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613) 237-6900
; TELEFAX: (613) 237-0045
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1016 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Human immunodeficiency virus type 1
; STRAIN: BRU
; US-07-743-357-5

Query Match 57.0%; Score 3109; DB 2; Length 1016;
Best Local Similarity 56.3%; Pred. No. 4.6e-257;
Matches 565; Conservative 177; Mismatches 242; Indels 20; Gaps 4;

QY 46 ARE---KTERARETRIQGSDRGLTAPRAGGDTIQGATNRLAA-----PQFS 89
Db 13 AREFSEQTRANSPTISSEQTRANSPTREQLQVWGRDNNLSLEAGDRQGTVSFNPQIT 72
QY 90 LMKRPVTVAYTEGQPVVLLDTGADDSIVAGTELNNYSPKIVGIGGIFGINTKEYKNVEI 149
Db 73 LMQRPVTVAYTEGQPVVLLDTGADDSIVAGTELNNYSPKIVGIGGIFGINTKEYKNVEI 132
QY 150 EVLNKVRATWGTPTNIFGRNLTALGMSLNLPVAKVEPIKMLPKGDKPKLRWMP 209
Db 133 EICGHAKGTVLVGTPTNIFGRNLTALGMSLNLPVAKVEPIKMLPKGDKPKLRWMP 192
QY 210 LTKKEIALKEICEKMEKEGQLEAPPTNPTPTFAIKKDKNKRMLIDFRELNVQ 269
Db 193 LTKKEIALKEICEKMEKEGQLEAPPTNPTPTFAIKKDKNKRMLIDFRELNVQ 252
QY 270 DFTETQLGIPHPAGLAKRRITVLDVGDAYSIPLHEDFRPYTAFTLPSVNNAPGKRYI 329
Db 253 DFWEVQLGIPHPAGLAKRRITVLDVGDAYSIPLHEDFRPYTAFTLPSVNNAPGKRYI 312
QY 330 YKVLPGQWKGSPAIQFHTMRQVLEPRFRKANKDVIIIOYMDLIIASDRDLEHDRVVLQ 389
Db 313 YKVLPGQWKGSPAIQFHTMRQVLEPRFRKANKDVIIIOYMDLIIASDRDLEHDRVVLQ 372
QY 390 KELLNGLGFSPTDEKQDPDPYHMGYELWPTKWKQLQIPLPQKEIWTVDIQLVGLVNW 449
Db 373 KELLNGLGFSPTDEKQDPDPYHMGYELWPTKWKQLQIPLPQKEIWTVDIQLVGLVNW 432
QY 450 WAAQYIPGKIKYKHLRGLRGKMTLFEVQWTELAELAEENRIILSQOEGHYOEEKELE 509
Db 433 WAAQYIPGKIKYKHLRGLRGKMTLFEVQWTELAELAEENRIILSQOEGHYOEEKELE 492
QY 510 EATYQKQOENQWYKIHQEE-KILKVGKAKYKNTHTNGIRLLAQVQKIGKEALVWGR 568
Db 493 EATYQKQOENQWYKIHQEE-KILKVGKAKYKNTHTNGIRLLAQVQKIGKEALVWGR 552
QY 569 IPKFLPVEREWEQWQWNYQWTVIPDWDVSTPPLVRLAFNLVGDPIPGAETTYTDS 628
```



```

; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Human immunodeficiency virus type 1
; STRAIN: HXB2
; US-07-743-357-22

Query Match          55.0%; Score 2999; DB 2; Length 913;
Best Local Similarity 58.6%; Pred. No. 1e-247;
Matches 535; Conservative 166; Mismatches 208; Indels 4; Gaps 2;

QY 121 IELGNNYSKPIVGGIGGFINTKEYKNVEIEVLNKKVVRATIMTGDTPINIFGRNLTALGM 180
Db 1 MSLPGRWKPKMIGGIGGFIVRQYDQILIEICGHKAIGTVLVGPTPVNIIGNRLLTQIGC 60

QY 181 SLNLPVAKVEPIKIMLPGKDGPKLRQWPLTKEIKALKEICEKMEKEGLEEAPTNPY 240
Db 61 TLNFPISPETVTVKLVKPGMDGPKVKQWPLTEKIKALVEICTEMEKEGKISKIGPENY 120

QY 241 NPTPTFAIKKDKKNKRWMLIDFRELNKVTQDFTFQIGIPHPAGLAKKRRTITVLDVGDAVF 300
Db 121 NTPVFAIKKDKSTKRWKLVDFRELNKARTQDFEVQLGIPHPAGLAKKKSVTVLDVGDAVF 180

QY 301 SIPLHEDFRPYTAFTLPSVYNNAEPPGKRYIYKVLPGQWKSIPAIFQHTMRQVLEPPRKANK 360
Db 181 SVPLDDEFRKYTAFTIPSIINNETPGIRYOYNVLPQGWKSPAIQFSSMTKILEPPRKQNP 240

QY 361 DVILIOYMDLIIASDRTDLEHDRVVLQKLKELLNGLGFSTPDEKFOKDPPIYHMGYELWP 420
Db 241 DIVIYOYMDLVGSDELQHTKIEELRQHLRWGLTTPDKKHQKEPFFLWMGYELHP 300

QY 421 TKWKLQIKLOPEKIWTVDIOKLVGLNNAQLPGIKTKHLCLRLIRGKMTLTERVQWT 480
Db 301 DKWTQPIVLPEKDSWTVDIOKLVGKLNWASQIPIGKVRQLCKLRTKALTEVPLT 360

QY 481 ELAEALEENRIILSQEGHYHQBKELEATVQKDQENQWYTKIHQEE-KILKVGKYAK 539
Db 361 EBAELEAENREILPEVHGVIYDPSKDIAETQKQCGQWYQIYQEPFKNLTKYAR 420

QY 540 VNKTHNGIRLLAQQVQKIGKALVTVGRIPKFLHPVERIEFWQWMDNTWQVTPDWF 599
Db 421 MKGAHTNDVKQTEAVQKITTISIVTGWTPKPKLPDKETWETWTWYQATWPEWBF 480

QY 600 VSTPPLVRLAFNLVGPDPGAETFTYDGCNRSQSKGAGYVTDGRKDKVKKLEQTTNQ 659
Db 481 VNTPLVLKWLQLEKPIVGAETFTYDGAANRETKLGKAGYVYTNRGKQKVVLTDTNOK 540

QY 660 AELEAFAMALTDGPKVNIIVDSQVYVGMISASOPTESKIVNQIIIEEMIKKEAIVAVY 719
Db 541 TELQATYALQDSGLEVNIVTDSQVAGLIQAQDQSESELVNQIIIEQLIKKEKVLAWY 600

QY 720 PAHKGIGGQNEVDHLVSQIGRQVLEKLETPAQEBHEKYHNSVKLSHKFGIPNLVAROI 779
Db 601 PAHKGIGGNEQVDKLVKSAGIRKVLFDGIDKAQDEHEKYHNSWRAMASFNLPVYVAKEI 660

QY 780 VNSCAQCGQGAIRHGVQNAELGTWQMDCTHLEGKIIIVAVHVASGFIAEVIPOBSGRQ 839
Db 661 VASCDKCOLKGAHMGQVDCSPGIMQLDCTHLEGKVLVAVHVASGYIEAEVIPAETGOE 720

QY 840 TALFLKLASRPWITHLTNDGANFTSQEVKVMVNWIGIEQSGFYVPYPSQSGVVEAMNH 899
Db 721 TAYFLKLKAGRPVKTITHTDNGSNFTGATVRAACWNAWIKQFEGIPYPSQSGVVEAMNK 780

QY 900 HLKNOISRIREQANTLETIVLMAIHCMMNFKRGGIGDMTPSERLNMWITTEQIEQLOAK 959
Db 781 ELUKTIQVRODQAEKLTAVQMAVFIHNFKRKGGIGGYSGAGRIVDIATDITQLELQK 840

QY 960 NSKLKDFRYFREGRDQLWKGPCELLWKGEGLVKGVDIKIIPRRKAIIRDYGGROE 1019
Db 841 ITKIONFRVYRDSRNSLWKGPALKLWKGEAGVAVIQDNSDIKVPVPRKAIIRDYKGOM- 899
```

```

QY 1020 MDSGSHLEGARED 1032
Db 900 --AGDDCVASRQD 910

RESULT 21
US-07-743-357-6
; Sequence 6, Application US/07743357
; Patent No. 5858646
; GENERAL INFORMATION:
; APPLICANT: Kang, Yong C.
; TITLE OF INVENTION: Polypeptide having immunological
; activity for use as diagnostic reagent and/or vaccine
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KIRBY EADES GALE BAKER
; STREET: Box 3432, Station D
; CITY: Ottawa
; STATE: Ontario
; COUNTRY: Canada
; ZIP: K1M 1H8
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/743,357
; FILING DATE: 21-AUG-1991
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/CA90/00062
; FILING DATE: 23-FEB-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Gale, Edwin J.
; REGISTRATION NUMBER: 28,584
; REFERENCE/DOCKET NUMBER: 30924-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613) 237-6900
; TELEFAX: (613) 237-0045
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 913 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FRAGMENT TYPE: Internal
; ORIGINAL SOURCE:
; ORGANISM: Human immunodeficiency virus type 1
; STRAIN: MN
; US-07-743-357-6

Query Match          54.2%; Score 2957; DB 2; Length 913;
Best Local Similarity 57.8%; Pred. No. 4.1e-244;
Matches 528; Conservative 166; Mismatches 215; Indels 4; Gaps 2;

QY 121 IELGNNYSKPIVGGIGGFINTKEYKNVEIEVLNKKVVRATIMTGDTPINIFGRNLTALGM 180
Db 1 MSLPGRWKPKMIGGIGGFIVRQYDQITIGICGHKAIGTVLVGPTPVNIIGNRLLTQIGC 60

QY 181 SLNLPVAKVEPIKIMLPGKDGPKLRQWPLTKEIKALKEICEKMEKEGLEEAPTNPY 240
Db 61 TLNFPISPETVTVKLVKPGMDGPKVKQWPLTEKIKALVEICTEMEKEGKISKIGPENY 120

QY 241 NPTPTFAIKKDKKNKRWMLIDFRELNKVTQDFTFQIGIPHPAGLAKKRRTITVLDVGDAVF 300
Db 121 NTPVFAIKKDKSTKRWKLVDFRELNKKTQDFEVQLGIPHPAGLAKKKSVTVLDVGDAVF 180

QY 301 SIPLHEDFRPYTAFTLPSVYNNAEPPGKRYIYKVLPGQWKSIPAIFQHTMRQVLEPPRKANK 360
Db 181 SVPLDDEFRKYTAFTIPSIINNETPGIRYOYNVLPQGWKSPAIQFSSMTKILEPPRKQNP 240
```



```
; APPLICANT: Barr, Margaret C.
; TITLE OF INVENTION: No. 6284253el Feline Immunodeficiency Virus Nucleotide Sequence
; FILE REFERENCE: 18617.0059
; CURRENT APPLICATION NUMBER: US/09/238,303B
; CURRENT FILING DATE: 1999-01-28
; EARLIER APPLICATION NUMBER: US 60/072,927
; EARLIER FILING DATE: 1998-01-29
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 9
; LENGTH: 1150
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: protein encoded by the pol gene of a recombinant viral
; OTHER INFORMATION: clone constructed from the genomic DNA of a Pallas's cat feline
; OTHER INFORMATION: immunodeficiency virus
US-09-238-303-9

Query Match      33.2%; Score 1810; DB 3; Length 1150;
Best Local Similarity 33.9%; Pred. No. 1.1e-145;
Matches 393; Conservative 222; Mismatches 347; Indels 198; Gaps 18;

QY      8 GPLGKEAPQLPRGPSSAGADTNTPSGSSSGSTGEIYAAREKTERAERETIQGSDRLTA 67
DB      13 GHAGKAAGKLAEGSCCPYQTSAAISNSSINNS-----ESATWSINTAFGSSNGVP- 63
QY      68 PRAGGDTIOGATNRGLAAPQFSLWKRPVVTAYIEGQPVVLLDTGAD-          115
DB      64 -----YGLKRD-----ELIHRPILIIYVNGPIRFLMDTGADITIMNAEDNILN 110
QY      116 SIVAGIELGNYSKPIVGGIGGFINTKEYKNVEIYLNKKVRATIMTGDTPI----- 167
DB      111 SIPDGIQ-----TWIGVGGKGRKFRVRHLEIRDPNHRQAQCLFGNMCILDDNSLTE 162
QY      168 NIFGRNLTALGMSLNP--VAKVEPIKMLKCGKDGPKLRQWPLFKEKEALKECEKM 225
DB      163 PLLGRONMVRFGAKLVANISNKIPIVKVMKDPKSGPKIKQWPLSKEKEALTEIVYRL 222
QY      226 EKEGQLEAEAPTPNPTPTFAIKKDKNKRMLIDFRELKNVTDTEIQLGPHPAGLA 285
DB      223 EKEGVKVRADPNPWTPIFCIKK-SGKWRMLIDFTLNELTEKAEVOLGLPHPAGLQ 281
QY      286 KKRITVLDVGDAFYSLPHEDFRPYTAFTPLPSVNNAPGKRIYKVLPGWKGSPAIFQ 345
DB      282 ERQQTVDLADAYFTIPLDPYAPYTAFTPLPKINNSGGERFVWCGLPQGWVSLPIYQ 341
QY      346 HTMQVLEPRKANKDVIIOYWDILIASDRTDLEHDRVVLQKELLNCLGFSTPDEKF 405
DB      342 STLNNILKPREQHEIDLYQYDDIYIGSDLGKKEHQIVEELRKLMLWGFETPEDKL 401
QY      406 QKDPPIYHMGYELWPTKWKQLQTLPOKEIWTVDIQLKVLNAAQYLPGLKTKHLCR 465
DB      402 QEOPPYKMWGYELPRKWTIQTRELIPEEPTLNELQKLVGIINWSSQIIPGLRIKALTN 461
QY      466 LIRGKMTLBEVOWTELAELAEENRIILSQOEGHYOEEKELEATVQKQDNQWYKI 525
DB      462 MMKGNOALDSKRWTEAKKEAEAKLAEIOHTQLGYDPQOQLHAKLSIVGPHCIYQV 521
QY      526 HQS---EKILKVKYAKVKNTHNGIRLLAQVVKIGKEALVWGRIPKPHLPSVERIE 582
DB      522 YQGSFDPKLWYGMKNRQKKAENTCDIALRIYKIREESIVRLGKEPIYEIPCSREWE 581
QY      583 QMWDNYQVQWTI----PDMDFVSTPPLVRLAFNLV-GDPIPGAETFTYDGSNQRQSKEG 637
DB      582 ---SNLINTPYLACPPQVEYIHAAIMIQRSLSMIKEEPIRGAETWIDGG-RKKQSAR 637
QY      638 AGVTVDRGDKVKKLETTNQQAELFAFAMALDPSGPKVINIYVDSQYVWGISASQPTSE 697
DB      638 AAYTDTKGWEVWQIE-GSNQRAEVAMALLMARSGGEEMNIVDSQYILNLRQKPDLM 696
QY      698 SKVNOILIEBMIKKALYVAVWPAHKGIGNQOEVHIL----- 734
DB      697 G-LWQELLEIEKVAIFIDWPGHKIPGNTVDNLQCTMMIISNGILDKGEEDAGYD 755
```

```
QY      735 -----VSQGIQVQL----- 743
DB      756 LLAEQDIHLMPGEVRIVPTGVRMLPKGHGMVVGKSSIAKQGLDLVGLGVDEYRGEIG 815
QY      744 -----FLEK 747
DB      816 VININLQKRSITLKEKQKVAQLIIIPCKHEELKAGEIELNSERGEKGYGSTGAFASWNN 875
QY      748 IEPAQEEHEKYHNVKELSHKFGIPNLVARQIVNSCAQCOQKGEAIGHQVNAELGTWQMD 807
DB      876 IEAEINHEKFSHPDEFRLTEFGLPKQVABEIRKCPCLCTVQGEQVMGKLKVGPGIWD 935
QY      808 CTLEKGIILIVAVHVASGETAEAVIPQESGRQTALFLKLKASRWPTIHLHTDNGANFTSQ 867
DB      936 CTLEKGIILIVAVNTESGYIWARIIPOETADMTVKYLLQLISEHHVTELOSDNGPNPNA 995
QY      868 EVKVMVAWIGIEQSGFVYPNPQSGVVEAMNHLKKNQISIRIQANTTETIVLMAHCMN 927
DB      996 KVEGMTGFLGIKHKGYPGNPQSOALVENTNRMLEKWKIKFRGEVTTLDAALALALYALN 1055
QY      928 FKRRGGIGDMTPSERLINMITTEQETQFLOAKNSKLKDFVYFREGRDQLWKPGELLWK 987
DB      1056 FKQGRIGRISPYELLIQOESDRIDYFSKIPANNIKNSWYIKRRDKEMKGTQVEYW 1115
QY      988 GEGAVLVK-VGTDIKIIPRR 1006
DB      1116 GQGAVLIKHPEHGYMLIPRR 1135

RESULT 26
US-09-946-239-9
; Sequence 9, Application US/09946239
; Patent No. 6579527
; GENERAL INFORMATION:
; APPLICANT: Barr, Margaret C.
; TITLE OF INVENTION: No. 6579527el Feline Immunodeficiency Virus Nucleotide and
; FILE REFERENCE: 18617.0059
; CURRENT APPLICATION NUMBER: US/09/946,239
; CURRENT FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: US 09/238,303, US 60/072,927
; PRIOR FILING DATE: 1999-01-28, 1998-01-29
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 9
; LENGTH: 1150
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: protein encoded by the pol gene of a recombinant viral
; OTHER INFORMATION: clone constructed from the genomic DNA of a Pallas's cat feline
; OTHER INFORMATION: immunodeficiency virus
US-09-946-239-9
```

```
Query Match      33.2%; Score 1810; DB 4; Length 1150;
Best Local Similarity 33.9%; Pred. No. 1.1e-145;
Matches 393; Conservative 222; Mismatches 347; Indels 198; Gaps 18;

QY      8 GPLGKEAPQLPRGPSSAGADTNTPSGSSSGSTGEIYAAREKTERAERETIQGSDRLTA 67
DB      13 GHAGKAAGKLAEGSCCPYQTSAAISNSSINNS-----ESATWSINTAFGSSNGVP- 63
QY      68 PRAGGDTIOGATNRGLAAPQFSLWKRPVVTAYIEGQPVVLLDTGAD-          115
DB      64 -----YGLKRD-----ELIHRPILIIYVNGPIRFLMDTGADITIMNAEDNILN 110
QY      116 SIVAGIELGNYSKPIVGGIGGFINTKEYKNVEIYLNKKVRATIMTGDTPI----- 167
DB      111 SIPDGIQ-----TWIGVGGKGRKFRVRHLEIRDPNHRQAQCLFGNMCILDDNSLTE 162
QY      168 NIFGRNLTALGMSLNP--VAKVEPIKMLKPGKDGPKLRQWPLFKEKEALKECEKM 225
DB      163 PLLGRONMVRFGAKLVANISNKIPIVKVMKDPKSGPKIKQWPLSKEKEALTEIVYRL 222
```


Db 1 QWPLTEKIKALVEICTEMEKEGKISKIGPENPNTPVFAIKKDKSTKWRKLVDFRELNK 60
Qy 267 VTQDFTEIQLGIPHPAGLAKRRRTITVLDVGDAYESIPLHEDFRPYTAFTLPSVNNAPGK 326
Db 61 RTQDFWEVQLGIPHPAGLAKRRRTITVLDVGDAYESIPLHEDFRPYTAFTLPSVNNAPGK 120
Qy 327 RYIKVLPQGWKSGPAIFQHTMRQVLEPFRKANKDVIIIOYMDLILASDRDLEHDRV 386
Db 121 RYQNVLPQGWKSGPAIFQSSMTKILEPFRKONPDIVIQYMDLILASDRDLEHDRV 180
Qy 387 LQKELLNGLGFSPTDEKFKQDPYHMGVYELWTKWKLQIKLQIPQKEIWTVDIQLV 446
Db 181 EELSHLLRLMGLTTPDKKHQKPEPFLMWGVELHDPKWTQVIVLPEKDSWTVNDIQLV 240
Qy 447 VLNMAALYPGIKTKHLRIRGKWLTEEVQWTEELAELEENRIILSOBQEGHYOEE 506
Db 241 KLNWASQIYPGIKVYRQLCKLLRGTKALTEVILPTEEALELAENREILKEPVHGVYDPS 300
Qy 507 KELEATVOK 515
Db 301 KDLIAEIQK 309

RESULT 32

US-08-987-867A-6
; Sequence 6, Application US/08987867A
; Patent No. 6063384
; GENERAL INFORMATION:
; APPLICANT: C. Morrow et al.
; TITLE OF INVENTION: ENCAPSULATED RECOMBINANT VIRAL
; TITLE OF INVENTION: NUCLEIC ACID AND METHODS OF MAKING AND
; TITLE OF INVENTION: USING SAME
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 STATE STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/987,867A
; FILING DATE: 09-DEC-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/087,009
; FILING DATE: 01-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: UAP-004CPDV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 742-4214
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 314 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-987-867A-6

Query Match 19.6%; Score 1068; DB 3; Length 314;
Best Local Similarity 63.4%; Pred. No. 4.4e-83;
Matches 196; Conservative 41; Mismatches 72; Indels 0; Gaps 0;

Qy 207 QWPLTEKIKALVEICTEMEKEGKISKIGPENPNTPVFAIKKDKKWRKLVDFRELNK 266

Db 1 QWPLTEKIKALVEICTEMEKEGKISKIGPENPNTPVFAIKKDKSTKWRKLVDFRELNK 60
Qy 267 VTQDFTEIQLGIPHPAGLAKRRRTITVLDVGDAYESIPLHEDFRPYTAFTLPSVNNAPGK 326
Db 61 RTQDFWEVQLGIPHPAGLAKRRRTITVLDVGDAYESIPLHEDFRPYTAFTLPSVNNAPGK 120
Qy 327 RYIKVLPQGWKSGPAIFQHTMRQVLEPFRKANKDVIIIOYMDLILASDRDLEHDRV 386
Db 121 RYQNVLPQGWKSGPAIFQSSMTKILEPFRKONPDIVIQYMDLILASDRDLEHDRV 180
Qy 387 LQKELLNGLGFSPTDEKFKQDPYHMGVYELWTKWKLQIKLQIPQKEIWTVDIQLV 446
Db 181 EELSHLLRLMGLTTPDKKHQKPEPFLMWGVELHDPKWTQVIVLPEKDSWTVNDIQLV 240
Qy 447 VLNMAALYPGIKTKHLRIRGKWLTEEVQWTEELAELEENRIILSOBQEGHYOEE 506
Db 241 KLNWASQIYPGIKVYRQLCKLLRGTKALTEVILPTEEALELAENREILKEPVHGVYDPS 300
Qy 507 KELEATVOK 515
Db 301 KDLIAEIQK 309

RESULT 33

US-08-679-493A-72
; Sequence 72, Application US/08679493A
; Patent No. 6303295
; GENERAL INFORMATION:
; APPLICANT: Taylor, Ethan W.
; TITLE OF INVENTION: SELENOPROTEINS, CODING SEQUENCES AND METHODS
; FILE REFERENCE: 55-95
; CURRENT APPLICATION NUMBER: US/08/679,493A
; CURRENT FILING DATE: 1996-07-12
; PRIOR APPLICATION NUMBER: 60/001203
; PRIOR FILING DATE: 1995-07-14
; PRIOR APPLICATION NUMBER: 60/003,112
; PRIOR FILING DATE: 1995-09-01
; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 72
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)..(327)
; OTHER INFORMATION: X is selenocysteine.
US-08-679-493A-72

Query Match 19.4%; Score 1055; DB 4; Length 327;
Best Local Similarity 59.5%; Pred. No. 6.1e-82;
Matches 198; Conservative 44; Mismatches 79; Indels 12; Gaps 2;

Qy 185 PVAKEPIKIMLPGKDGPKLRWMLTKERIEALKEICEKMEKEGQLEEEAPPNTPT 244
Db 1 PISPITVPVKLPKPGMDGPKVQWPLTEEEKIKALVEICTEMEKEGKISKIGPENYTPV 60
Qy 245 FAIKKDKKWRKLVDFRELNKVTQDFTEIQLGIPHPAGLAKRRRTITVLDVGDAFSIPL 304
Db 61 FAIKKDKSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLAKRRRTITVLDVGDAFSIPL 120
Qy 305 HEDFRPYTAFTLPSVNNAPGKRYIKVLPQGWKSGPAIFQHTMRQVLEPFRKANKDVII 364
Db 121 DEDFRPYTAFTLPSVNNAPGKRYIKVLPQGWKSGPAIFQSSMTKILEPFRKONPRHSY 180
Qy 365 IQYMDLILASDRDLEHDRVVLOKELLNGLGFSPTDEKFKQDPYHMGVYEL 418
Db 181 LSTHGXFCVCRIXLEIGHRTKIE-----ELRQHLLRWGLTTPDKKHQKPEPFLMWGVEL 234
Qy 419 WPTKWLKLOLQIPQKEIWTVDIQLVGLNMAALYPGIKTKHLRIRGKWLTEEVQ 478
Db 235 HPDKWTQVIVLPEKDSWTVNDIQLVGLNMAALYPGIKTKHLRIRGKWLTEEVQ 294


```

QY 479 WTELAEELEENRIILSOEQEGHYOEKELEA 511
DB 295 LTERAELEAENREILKEPVHGVIYDPSKDLIA 327

RESULT 34
US-08-679-493A-71
; Sequence 71, Application US/08679493A
; Patent No. 6303295
; GENERAL INFORMATION:
; APPLICANT: Taylor, Ethan W.
; TITLE OF INVENTION: SELENOPROTEINS, CODING SEQUENCES AND METHODS
; FILE REFERENCE: 55-95
; CURRENT APPLICATION NUMBER: US/08/679,493A
; CURRENT FILING DATE: 1996-07-12
; PRIOR APPLICATION NUMBER: 60/001203
; PRIOR FILING DATE: 1995-07-14
; PRIOR APPLICATION NUMBER: 60/003,112
; PRIOR FILING DATE: 1995-09-01
; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 71
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
; NAME/KEY: VARIANT
; LOCATION: (1)..(327)
; OTHER INFORMATION: X is selenocysteine.
US-08-679-493A-71

Query Match 19.2%; Score 1048; DB 4; Length 327;
Best Local Similarity 59.0%; Pred. No. 2.4e-81;
Matches 193; Conservative 46; Mismatches 86; Indels 0; Gaps 0;

QY 185 PVAKVEPIKIMLKPGDKPKLRQWPLTKKEIKAEKCEKMEKEGQLEAEPTNPYPT 244
DB 1 PISPETVPVKLKPGMDGPKVKQWPLTEEEKIKALVEICTEMEKEGKISKIGPENYP 60

QY 245 FAIKKKNKRWMLIDFRELNVQDFTETIQLGIPHPAGLAKKRITVLDVGDYFSL 304
DB 61 FAIKKDKSTRKRVLDRELNKRQDQFEWVQLGIPHPAGLAKKRISNSTGCGXCIF 120

QY 305 HEDFRPYTAFTLPSVNAEPCKRYIYKVLPGWKGSPAIFQHTMRQVLEPPRKANKOV 364
DB 121 DEDFRKYTAFTIPINNTPGIRYQYVNLPGWKGSPAIFQSSMTKILEPFPKQNPDI 180

QY 365 IQYMDDLIASDRTDLEHDRVVLQKELLNGLGFSTPDEKFKDPPYHWMGYELMPTKW 424
DB 181 YOYMDLVVSGDLEIGQHRKIEELRQHLRWGLTTPDKKHQKEPPFLMGMGYELHP 240

QY 425 LQKIQLOKETEWTNDIQKLVGLNWAQAQLPGIKTKHLIRGKWTLTTEEVQWTELA 484
DB 241 VQIYVLPKESWTNDIQKLVGLNWAQAQLPGIKTKHLIRGKWTLTTEEVQWTELA 300

QY 485 AELEENRIILSOEQEGHYOEKELEA 511
DB 301 LELAENREILKEPVHGVIYDPSKDLIA 327

RESULT 35
5320958-3
; Patent No. 5320958
; APPLICANT: INOUE, SUNIKO; HSU, MEI-YIN; EAGLE, SUSAN;
; INOUE, MASAYORI
; TITLE OF INVENTION: ISOLATED BACTERIAL REVERSE TRANSCRIPTASE
; NUMBER OF SEQUENCES: 24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/315,316
; FILING DATE: 24-FEB-1989
; SEQ ID NO:3
; LENGTH: 261

Query Match 17.4%; Score 948; DB 6; Length 261;
Best Local Similarity 65.0%; Pred. No. 6.1e-73;
Matches 173; Conservative 37; Mismatches 48; Indels 8; Gaps 3;

QY 194 IMLKPGDKPKLRQWPLTKKEIKAEKCEKMEKEGQLEAEPTNPYPTFAIKKDKKN 253
DB 1 VKLKPGMDGPKVKQWPLTEEEKIKALVEICTEMEKEGKISKIGPENYPVFAIKKDK 60

QY 254 KRWMLIDFRELNVQDFTETIQLGIPHPAGLAKKRITVLDVGDYFSLPLHEDFRPY 313
DB 61 KRWKLVDFRELNKRQDQFEWVQLGIPHPAGLAKKRISVTVLDVGDYFSLPLDED 119

QY 314 FTLPVNAEPCKRYIYKVLPGWKGSPAIFQHTMRQVLEPPRKANKOVIIQYMDDL 373
DB 120 FTIPSINNETPGIRYQYVNLPGWKGSPAIFQSSMTKILEPFPKQNPDIYQYMD 179

QY 374 ASDRTDLE---HDRVVLQKELLNGLGFSTPDEKFKDPPYHWMGYELMPTKWIK 430
DB 180 G----DLEIGQHRKIEELRQHLRWGLTTPDKKHQKEPPFLMGMGYELHPDKW 235

QY 431 POKEIWTVDIQKLVGLNWAQAQLYP 456
DB 236 PEKDSWTVDIQKLVGLNWAQAQLYP 261

RESULT 36
US-09-231-182B-2
; Sequence 2, Application US/09231182B
; Patent No. 6303334
; GENERAL INFORMATION:
; APPLICANT: HOLLER, TOD PAUL
; APPLICANT: MEYER, ANNETTE
; APPLICANT: NABEL, GARY JAN
; APPLICANT: POST, LEONARD
; TITLE OF INVENTION: METHOD OF USING TRANSDOMINANT NEGATIVE RETROVIRAL
; FILE REFERENCE: 4105.53USD2
; CURRENT APPLICATION NUMBER: US/09/231,182B
; CURRENT FILING DATE: 1999-01-14
; PRIOR APPLICATION NUMBER: 08/841,179
; PRIOR FILING DATE: 1997-04-29
; PRIOR APPLICATION NUMBER: 08/286,578
; PRIOR FILING DATE: 1994-08-05
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: HIV Integrase
US-09-231-182B-2

Query Match 17.0%; Score 924.5; DB 4; Length 290;
Best Local Similarity 57.1%; Pred. No. 7.5e-71;
Matches 165; Conservative 60; Mismatches 61; Indels 3; Gaps 1;

QY 744 FLEKIEPAQEEHEKYHSHNVKELSHKFGIPNLVARQIVNSCAQCQKQGEAIGHQVNAELGT 803
DB 3 FLGDIDKAQEEHEKYHSHNWRAMASDFNLPPVVAKEIVASCQKQKGEAMHGQVDCSPGI 62

QY 804 WQMDCTHLEGKIIIVAVHVASGFIAEVIPOESGRQATFALLKLASRPITHLHNDGAN 863
DB 63 WQMDCTHLEGKIIIVAVHVASGFIAEVIPOESGRQATFALLKLASRPITHLHNDGAN 122

QY 864 FTSEVKNVAVWIGIEQSGFVPYNPQSGVYEAAMHHLKNOISRRECAQNTIETVLNAI 923
DB 123 FTSITVKAACWAGIKQEGFIPYNPQSGVYEAAMHHLKNOISRRECAQNTIETVLNAI 182

QY 924 HCMNFKRRGGIGDMTPPSERLINMITTEQEIQFQAKNSKLKDFRYVYFREGDQLKGPGE 983
; SEQ ID NO:3
; LENGTH: 261
```



```
US-07-648-796A-6
; Sequence 6, Application US/07648796A
; Patent No. 5310876
; GENERAL INFORMATION:
; APPLICANT: Dr. Hubert Bayer
; APPLICANT: Dr. Erhard Kopetzki
; TITLE OF INVENTION: Expression of HIV1 and HIV2
; TITLE OF INVENTION: Polypeptides and their use
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/648,796A
; FILING DATE: 19910125
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 40 02 636.1
; FILING DATE: 30-JANUARY-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5310876man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: HUBR 784
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 376 amino acid residues
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-648-796A-6

Query Match 15.2%; Score 827.5; DB 1; Length 376;
Best Local Similarity 54.2%; Pred. No. 2.4e-62;
Matches 156; Conservative 56; Mismatches 59; Indels 17; Gaps 4;

QY 740 ROVLFE-----KIPAEHEKYSNVKELS--HKFGIPNLVARQIVNSCAQCOQKGEA 792
DB 86 KQVRYLEARNISKLSQAQIQEK--NMVELQKLSNWDPP-----LSCDKCOLKGEA 135

QY 793 IHGOVNAELGTWQMDCTHLEGLIIVAVHVASGFTEAEVPOESGROTALFLKLASRP 852
DB 136 MHGQVDCSPGIWQDCTHLEGLIIVAVHVASGFTEAEVPAETGOETAYFIKLGRWP 195

QY 853 ITHLHTDNGANFTSEVKVAVWIGIEQSGVYPNQSQGVVYVAMNHLKNOISIREQA 912
DB 196 KVIHTDNGSNFTSTVRAACWAGIKQEFPIPNPQSQGVVYVAMNHLKNOISIREQA 255

QY 913 NTIETVLMIAHMCNKRGGDMTPSERLINMITTEQEIQLQAKNSKLKDFRYFRE 972
DB 256 EHLKTAQVQMAVFIHFKRGGIGGYGYSAGERIYDIATDIQTKELQKIIKIQNFYRYRD 315

QY 973 GRDQLWKGPGELLWKEGAVLVKGTDIKIIPRRKAKIIRDYGGQREM 1020
DB 316 SRDPLWKGPAKLLWKEGAVVQDNSEIKVPRRKAIRDYGKQSDL 363

RESULT 40
5252477-3
; Patent No. 5252477
; APPLICANT: OROSZIAN, STEPHEN; COPELAND, TERRY D.
; TITLE OF INVENTION: HUMAN IMMUNODEFICIENCY VIRUS SPECIFIC

; PROTEOLYTIC ENZYME AND A METHOD FOR ITS SYNTHESIS AND RENATURATION
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/57,183
; FILING DATE: 01-JUN-1987
; SEQ ID NO: 3:
; LENGTH: 275
5252477-3

Query Match 15.1%; Score 821; DB 6; Length 275;
Best Local Similarity 56.2%; Pred. No. 5e-62;
Matches 150; Conservative 43; Mismatches 72; Indels 2; Gaps 1;

QY 55 RETIQGSDRGLTAPRAGGTIOCATNRGLAARQFSLWKRPPVVTAYTEGQVPEVLLDTGAD 114
DB 9 RRELQVWGRDNNSPSEAGADROGTYSFNF--PQITLWQRPVLTIKIGGLKEALLDTGAD 66

QY 115 DSIVAGIELGNYSKPIGVGIGGFINTKYNKVEIEVLNKKVRATIMTGDTPINIFGRNI 174
DB 67 DTVLEMSLPGRWKPKMIGGIGGFIVROYDQILIEICGHKAIGTVLVGPTPVNIIGRNL 126

QY 175 LTALGMSLNLPAKVEPIKIMLKPGKDGPKLRQWPLTKKIEALKEICEKMEKEGOLEEA 234
DB 127 LTQIGCTLNFPIETVVPVKLPKPGMDGPKVQWPLTEELIKALVEICTEMEKEGKISKI 186

QY 235 PPTNPYNTPTFAIKKKDKNKRMLIDFRELNKVTDFTETIQLGIPHPAGLAKRRITVLD 294
DB 187 GPENPYNTPVFAIKKKDKSTKWRKLVDFRELNRKTQDFWEVQLGIPHPAGLAKKKSVTLD 246

QY 295 VGDVFSIPLHEDFRPYTAFTLPSVNN 321
DB 247 VGDVFSVPLDEDFRKYTAFTIPSINN 273

RESULT 41
US-07-648-796A-2
; Sequence 2, Application US/07648796A
; Patent No. 5310876
; GENERAL INFORMATION:
; APPLICANT: Dr. Hubert Bayer
; APPLICANT: Dr. Erhard Kopetzki
; TITLE OF INVENTION: Expression of HIV1 and HIV2
; TITLE OF INVENTION: Polypeptides and their use
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/648,796A
; FILING DATE: 19910125
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 40 02 636.1
; FILING DATE: 30-JANUARY-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5310876man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: HUBR 784
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 247 amino acid residues
```

```

; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-648-796A-2

Query Match      14.7%; Score 804; DB 1; Length 247;
Best Local Similarity 60.9%; Pred. No. 1.2e-60;
Matches 142; Conservative 45; Mismatches 46; Indels 0; Gaps 0;

Qy 783 CAQCOQGEATHGQVNAELGTWQMDCHLEKIIIVAVHVHVSAGFIEAIVPQESGRTAL 842
Db 8 CDKCOLGAEAMHGQVDCSPGWLQDCHLEKIIIVAVHVHVSAGFIEAIVPQESGRTAY 67
Qy 843 FLLKASRPITLHTDNGANFTSQEVKMAVWVIGIEQSGFVYVYVQSGVVEAMNHLK 902
Db 68 FILKLAGRPVKVLIHTDNGSNFTSTYKAACWAGIKQEGIPYVQSGVWESMKNELK 127
Qy 903 NOISRIREQANTITVILMAHNCNFKRGGIDMTPTSERLINMITTEQIEQIFQAKNSK 962
Db 128 KIIGQVRDQAEHLKTAQVMVFIHFKRKGIGGYSAGERIVDIATDIOTKELQKQIK 187
Qy 963 LKDFRVFRGRDOLWKGPCLELLWKGAVLVKVGTDIKIIPRKAKIIRDYG 1015
Db 188 IQNFRVYRSDRLPWKGPAKLWKGEGAVVYQDNSEIKVYVPRKAKIIRDYG 240

```

```

RESULT 42
US-08-679-493A-73
; Sequence 73, Application US/08679493A
; Patent No. 6303295
; GENERAL INFORMATION:
; APPLICANT: Taylor, Ethan W.
; TITLE OF INVENTION: SELENOPROTEINS, CODING SEQUENCES AND METHODS
; FILE REFERENCE: 55-95
; CURRENT APPLICATION NUMBER: US/08/679,493A
; CURRENT FILING DATE: 1996-07-12
; PRIOR APPLICATION NUMBER: 60/001203
; PRIOR FILING DATE: 1995-07-14
; PRIOR APPLICATION NUMBER: 60/003,112
; PRIOR FILING DATE: 1995-09-01
; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 73
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (199)
; OTHER INFORMATION: X is Q or Y.
; NAME/KEY: VARIANT
; LOCATION: (1)..(237)
; OTHER INFORMATION: X at positions 186, 192, 237 and 233 is
; OTHER INFORMATION: selenocysteine.
US-08-679-493A-73

```

```

Query Match      12.7%; Score 691; DB 4; Length 237;
Best Local Similarity 70.5%; Pred. No. 5.3e-51;
Matches 122; Conservative 23; Mismatches 28; Indels 0; Gaps 0;

Qy 185 PVAKVEPIKMLPGKDGPKLRWPLTKKIEALKEICEKMEKEGQLEAPPTNPYNTPT 244
Db 1 PISPITVPVVKLPGMDGPKVQWPLTEERIKALVEICTEMEKEGKISKIGPENPNTPV 60
Qy 245 FAIKKDKNKRMLIDPRELNKYTDPTETQLGPHIPAGLAKRRITVLDVGDAYFSIPL 304
Db 61 FAIKKDKSTKRWKLVDPRELNKYTDQDFEWQLGPHIPAGLAKRRITVLDVGDAYFSVPL 120
Qy 305 HEDFRPYTAFTLPSVNAEPGKRYIYKVLPGQWKGSPFAIQHTMROVLEPFRK 357
Db 121 DEDFRKYTAFTIPSINNETGIRYQYNLPGQWKGSPFAIQSSMTKILEPFRK 173

```

```

RESULT 43
US-08-259-451-11
; Sequence 11, Application US/08259451
; Patent No. 6406841
; GENERAL INFORMATION:
; APPLICANT: Lee, Helen H.
; APPLICANT: Swanson, Priscilla A.
; APPLICANT: Idler, Kenneth B.
; APPLICANT: Rosenblatt, Joseph D.
; APPLICANT: Chen, Irvin S. Y.
; APPLICANT: Golde, David W.
; APPLICANT: Robertson, Eugene F.
; APPLICANT: Stephens, John E.
; APPLICANT: Chan, Emerson W.
; APPLICANT: Buytendorp, Mark H.
; APPLICANT: Johnson, Joan E.
; APPLICANT: Motley, Cheryl T.
; APPLICANT: Peterson, Bryan
; APPLICANT: Edwards, Michelle
; APPLICANT: Guidinger, Peggy
; APPLICANT: Tate, Cynthia
; TITLE OF INVENTION: HTLV-IIINRA Compositions
; TITLE OF INVENTION: and Assays for Detecting HTLV Infection
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: One Abbott Park Road
; CITY: Abbott Park
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60064
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; MEDIUM TYPE: storage
; COMPUTER: IBM
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/259,451
; FILING DATE: 20-JUN-1994
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/086,415
; FILING DATE: 01-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Daniel W. Collins
; REGISTRATION NUMBER: 31,912
; REFERENCE/DOCKET NUMBER: 5381.US.PI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (708) 937-6365
; TELEFAX: (708) 938-2623
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 917 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
US-08-259-451-11

```

```

Query Match      10.8%; Score 588.5; DB 4; Length 917;
Best Local Similarity 25.1%; Pred. No. 3.2e-41;
Matches 233; Conservative 144; Mismatches 364; Indels 188; Gaps 39;

Qy 197 KPCKDGPKLROWPLT--KEKIEALKEICEKMEKEGQLEAPPTNPYNTPTFAIKKDKNK 254
Db 37 RPTPTPTAGPISPKPERKQALNDLVSKALEAGHIE--PYSGGNNPVPVKKN-GK 93
Qy 255 WRMLIDPRELNKYT-----QDFTQLGPHIPAGLAKRRITVLDVGDAYFSIPL 304
Db 94 WRHIDLRLATNATITLSPSPGPPDLTSLPTALPH-----LQTDLTDAFTQIPL 144
Qy 305 HEDFRPYTAFTLPSVNAEPGKRYIYKVLPGQWKGSPFAIQHTMROVLEPFRKANKDVII 364

```

Db 145 PROQFYFAFTIPQPCNYGEGTRYAWTVLPQGEKNSPTLFEQQLAAVLANPMRKMFTSTI 204
Qy 365 IOYMDILLASDRDLEHVRVVLQKLKELNGLGFSPTDEKFKQDP-----PYHW 413
Db 205 VOYMDILLASP-TNKLQOLSLTQALTTHGLPISEKQTPTPGQIRFLGOVISPNNH- 262
Qy 414 MGVELMPTKWKLOKIQLPQKEIWTWVNDIOKLVGLVNWAAQLYPGIKTKHLCLRI----- 467
Db 263 ITVESTPA-----IPIKSQWTLTELQVILGEIQWVSKGTP-ILRKHLSLYSALHGY 313
Qy 468 ---RGKWTLT-ERVQWTELAARELEN---RI-----ILSQEGHYQEKELEAT 512
Db 314 ROPRACITLTPQOLHALHAIQALQHNCRGRODPTPLGLGLSLSTSG-----TTSV 365
Qy 513 VOKDQEN---QWTYKIQHEKILKGVKAKVKNTHNGIRLLAQVVOVKIEKALVIWGR 569
Db 366 IFQPKQWPLAWLHTPHPTSLCPWG-----HLLACTILTLDKYTLQHYGLL 412
Qy 570 -PKFHLPVREIWEQWMDNY-----WQ-VTWIPDMDVFST 602
Db 413 CQSFHNMKSQALCDPLRNSPHSPVSGILIHMGREHNLGSPSPGWKTLHLHP-LLQ 470
Qy 603 PPLVRLAFNLVGDPIGATFTYDGCNRSQSKGAGYV---TDRGDKVKKLEQTN- 657
Db 471 PRLRPIFTLSPVLDTPACPLFSDGS-----PKAAYVLWDQTLQOQDITPLPPHETNS 524
Qy 658 -QOAELEA--FAMALTDSPGVNIIVDQYV-----MGISASQPTSESKIVN 702
Db 525 AQKGLLALIYGLRAAKPWSLNFILDSKYLKYLHSLAIGAPLGTSAHOTLOA----- 578
Qy 703 QIEEMIKKEAIYVAVPAHKIGG-----NOEVDHLVSQGIHQVLEKIEPAQEHEK 757
Db 579 -ALPPLLOGKTVLHVHRSHTNLPDPISFTNEYTDSLI-----VAPLVPLTP-QGLHGL 630
Qy 758 YHSNWKELSHKGIPLNVARQIVNSCAQOQGEATH---GQVNAEL---GTWQMDCTHL 811
Db 631 THCNQALY-SFGATPKAKSLVQTCQTIINSQHMPQGHIRGLLPNHNHMQGDVTHY 689
Qy 812 EKKIIVAVHV---ASGFIEAEVPOESGROT-ALFLKLASRWPIHLHTDNGANETS 866
Db 690 KYKKYKYLHVWVDFSGAVSVCKKETSCTISAFLOAISLLGLPHLINTDNGAPFLS 749
Qy 867 QEVKVMWVIGIEQSGVPYNPQSGVVEAMNHLKNOISR-IREQAN-TIETIVLMAIH 924
Db 750 QEFQECTSYHIKHSHPYNPNTSSGLVERTNGIKNLLNLYLLDCPNLPDINAINKALW 809
Qy 925 CMNFRRGIGDWTSPERLINMTTEQEIQFLOAKNSKLDFRYVREG-----RDQLWKG 980
Db 810 TLN-----QLNVNPNPQSGKTRWQIHHSPPPLPPEASTPPKPPSKWFFYKPLGLTNQWKG 864
Qy 981 PGLLWKGEGAVLVKVGTDIKIIPRRKAK 1009
Db 865 PLSLQEAAGAALLSIDGPRWIPWRFLK 893

RESULT 44

US-09-075-272-4
; Sequence 4, Application US/09075272
; Patent No. 6136598
; GENERAL INFORMATION:
; APPLICANT: MILLER, A. DUSTY
; APPLICANT: WOLGAMOT, GREG
; APPLICANT: BONHAM, LYNN
; TITLE OF INVENTION: MUS DUNNI ENDOGENOUS RETROVIRAL
; TITLE OF INVENTION: PACKAGING CELL LINES
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESS: TOWNSEND and TOWNSEND and CREW LLP
; CITY: San Francisco
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94111

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/075,272
; FILING DATE: 08-MAY-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/046,140
; FILING DATE: 09-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Poor, Brian W.
; REGISTRATION NUMBER: 32,928
; REFERENCE/DOCKET NUMBER: 14538A-003710
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1203 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-075-272-4

Query Match 10.6%; Score 576.5; DB 3; Length 1203;
Best Local Similarity 22.0%; Pred. No. 5.4e-40;
Matches 258; Conservative 190; Mismatches 413; Indels 313; Gaps 45;

Qy 94 PVVTAYIEGQPEVLELDTGADDSIVAGIELGNNSPK--IVGGIGGFI---NTKEYKNVE 148
Db 11 PVTLSVEGTPVNFLLDTGAESHVLS-PLGLGSRKTIVVGATGSKLIPWTKRALQID 69
Qy 149 IEVLNKKVRATIMTGDTPINIFGNILTAL----- 178
Db 70 ---KNMVTSHFLVPECPAPLLGRDLLTKLKAQVQFTSEGPQVSWGKAPLACLIVLSTEE 126
Qy 179 -----GMSLNLPVAKVEPKIMLPGKDGPKLQWPL 210
Db 127 YRLHEBPQGAAPLDWVTAPPNNVABEQAGMGL---AKQPPVNVVELKADATPISVRQYPM 183
Qy 211 TREKIEALKEICEKMEKQLEAPPTNPYPTFAIKKKKNKWRMLIDFRELNKVTD 270
Db 184 SKEAKGICPHIRLLDQGLIYAC--QSPWNPPLPVKPGTNDYRPVQDLREVNKRVL 241
Qy 271 FTEIQIGIPHPAGL-----AKKRITVLVDGDAYFISIPLEDPRPYTATFLPSVNNAPG 325
Db 242 ---IHTVTPNPYNLLSSLPPTWYTVLDLKDAFFCLRLHPKSQLLFAFEWRDPEGQGTG 298
Qy 326 KRYIYKVLPGQWKGSPAIQHTMRQVLEPFRANKDVIIIOYMDILLIASDRDLEHVR 385
Db 299 -QLTWTRLPQGFKNSTPDEALHRLADLAPFRAQNPQLTLQYVDDLLIAAASKELCQQT 357
Qy 386 VLQKELLNGLGFSPTDEKFO-KDPPYHMMGYEL-----WPTKWKLQTLQPKKEITVNV 439
Db 358 ERLTELGN-LGYRYSAKKAQICQTEVYILGTLRGCKRWLFEARKKTYWMLPPPT-TPR 415
Qy 440 DIQKLVGLNWAQAQLYPGIKTKHLCLIRGKMTLTEE---VOMTE----- 481
Db 416 QVREFLGTAGFCRLWIPGFAT-----LAAPLYPLTREGIPFEWKEHQFAEIKSSLMT 470
Qy 482 ---LAEAELEENRIILSQEGHYQEKELEATVQKDQENQWTKYKIHQEEKILKGVKA 538
Db 471 APALAPDLTKSFVL-----YVDERAGIARGVLTQALGPKWRPVAYLSK--KLDPPVA 520
Qy 539 KVKNTHTNGIRLLAQVVO-----KIGKEALVTWGRIPKPHLPVEREIEWQWMDN-----YV 589
Db 521 SCWPTCLKAIAAVALLIKDADKLTMGQQVTVV---APHALESIVRQPPDRWNTARNMTHY 577
Qy 590 QVTWIPDMDVFSTPPLVRLAF-----NLVGDPIPGAETFF 623

```
Db 578 QSLLLNDVTFAPPAILNPATLLPLTNDSDVPVHRCADILAEIEIGTRKDLTDQWPWGPAPSW 637
Qy 624 YTDGSCNRQSKGAGVYTDGDKV---KKLEOTTNQAELEAFAMALTDS-GPKVNI 679
Db 638 YTDGSSFLIEGKRAGAAVDGKVIWASALPEGSTSAQKAELIALTOALREAEKIINIY 697
Qy 680 VDSOYVMG-----ISASQPTSESKIVNQIIEEMIKKEAIIYVAVWPAH-K 723
Db 698 TDSRYAFATAHIHGAIYRQGLLTSGAKDIKNKEIL-ALLEAIHAPKKVAILHCPGHQK 756
Qy 724 G-----IGNGNEVDHL---VSQG-----TRQV-----LPLEK 747
Db 757 GEDLVAKGNRMADSVAKOVAQAGAMILTEKGNPKSPEDENYDIKELEFWTSDPLPYFPEGK 816
Qy 748 IEPAQOEHEKSHVNVKLSH-----KFGIPNL--VARQIVNSCAOC----- 786
Db 817 IDLTPEGIFKVGHLGHTHLGVEKMMRLIKKSRYQVFNLSKVAOKLIINSCKACAFNAT 876
Qy 787 -----QKGE--AIHQGVNAELGTWQMDCTHLE-----KIIIVAVHVSFGIFAEV 831
Db 877 KTYKEPGKRGQRGPGVY-----WEVDTEVKPMYGNKYLLVFDVTFSGWVEAF 927
Qy 832 IPQSGRQATALL-KLASRWPIH-LHTDNGANFTSQEVKVMVNWIGIEQSGVGPYNPQ 889
Db 928 TKTETAQIVAKKIFEEILPRYGVKVIQSDNGPAFVAQVSGGLATQGLIDMKLHCAYRPQ 987
Qy 890 SOGVVEAMNHLKNOISRIEQANTIETIVLMAI---HGMNFKRRGGI----- 934
Db 988 SSGQVERMNTLKETLKLAMEYGGKDWVALLPLAFRANTPGRFGLTFEVLVGGPPP 1047
Qy 935 -----GDMTPSERLINM-----ITTEQEIOTLOAKNSKLKDFRVYPREG 973
Db 1048 LIKDGTLVDPGSGVLPSSLLIHLKALKVIRTQIWQDLKATYPTGTTAVPHEFQV--- 1102
Qy 974 RDQLW-----KPGCELLWKGEVAVLVKVGVDIKI 1002
Db 1103 GDQVLRHRTGSLERWKGPYLLVLTPTAVKV 1136

RESULT 45
US-09-309-572-13
; Sequence 13, Application US/09309572
; Patent No. 6440730
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral hybrid vectors pseudotyped with LCMV
; FILE REFERENCE: P50489
; CURRENT APPLICATION NUMBER: US/09/309,572
; CURRENT FILING DATE: 1999-05-11
; EARLIER APPLICATION NUMBER: DE 198 56 463
; EARLIER FILING DATE: 1998-11-26
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 1737
; TYPE: PRP
; ORGANISM: Moloney murine leukemia virus
; FEATURE:
; OTHER INFORMATION: gag-pol protein
US-09-309-572-13

Query Match 10.2%; Score 554.5; DB 4; Length 1737;
Best Local Similarity 21.6%; Pred. No. 7.7e-38;
Matches 291; Conservative 193; Mismatches 466; Indels 397; Gaps 56;

Qy 18 PRGPSSAGADTN-----STPS-----GSSSGSTGETIYAAR-----EK 49
Db 381 PYDPEDPGQETINSMSPISQAPDIGRKLRLERLEDKKNKTLGDLVREAEIFNKRETPER 440
Qy 50 TERAERTIOGSDR-----GLTAPRAGGDTIOGATN 80
Db 441 EERIRTEBEKEARRTEDEQKEKERRRRHREMSKLLATVVSQKODRGGERRRSQLD 500
```

```
Qy 81 RLGA-----POFSLW-----KRPVVTAYIEGPV 105
Db 501 RDQCAVCKEKGHWAKDCPKPRGPRGPRPOTSLTLDDGGQGEPPPEPRITLKVGGQPV 560
Qy 106 EVLLDTGADDSIVAGIELGNYSK-----IVGGIGGFINTKEYK---NVEIEVLNKKV 156
Db 561 TFLVDTGAAQHSV-----LTQNPGLSDKSAAWQCATGG---KRYRTTDRKRVHLATGKV 611
Qy 157 RATIM-TGDRPPIINFRNILTALG-----MSLNL----- 184
Db 612 THSFLHVPDCPYPLGRLDRLTLKLAQIHFECSGAQVMPGQPLQVLTNLIEDEHRLHET 671
Qy 185 -----PVAKVE-----PIKMLKPGKDGPKLRQPLMTKEKIEA 217
Db 672 SKEPDVSLGSLTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIIKOYPNSQEARLG 731
Qy 218 LKEICEKMEKEGOLAEAPPTNPTPTFAIKKDKKWRMLIDFERELNKKVQTFETQLG 277
Db 732 IKPHIQRLDQGIL--VPCOSPWNTPLLPVKKPGTNDYRPVQDLREYNKREVD---IHPT 786
Qy 278 IHP-----AGLAKRR-ITVLDVGDVAFSIPLHEDERPYTAFTLPSVNNAEPCGRKIYK 332
Db 787 VNPYNLLSGLPSSHQYTVLDLKDFAFFCLRLHPTSQPLFAFEWRDEMGISG-QLTWTR 845
Qy 333 LPQGWKGSPAIFQHTMRQVLEPFRKANKDVIIQYMDIILIASDRTDLEHDRVVLQKEL 392
Db 846 LPOGFKNSTPLFDEALHRLADFRIOHPDLILILQYVDDLAA--TSEDCQOQTRALLQ 904
Qy 393 LNLGISTPDEK---FQKDPYHMGYEL-----WPTKKLQKIQLPQKEIWTVNDIQKL 444
Db 905 LGNLGYRASAKKAIQCKQVKY--LGYLLKEGORWLTAREKETV-MGQPTPKTPROREF 961
Qy 445 VGVNL--W-----AAOLYPIKTKHLCLIRGKMTLTERVQWT---ELAEAELE 488
Db 962 LGTAGFRLWIPGFAEWAAPLYPLTKGTLFNNGPQOQKAYQIKQALLTAPALGPLDT 1021
Qy 489 ENRIILSQEGBHYOEEKELEATVQKDOENQWYKIHOBKILKVGKVKAKVKNTHNGI 548
Db 1022 KPPELFVDEKQY-----AKGVLTQKLGWRRRPVAVLSK--KLDPVAAAGPPCLRMV 1071
Qy 549 RLQAQVVKIGK---EALVIGWRIPKFLPVREIWEQWMDN---YQVWTVPWDFV 600
Db 1072 AAIAVLTKDAGKLTMGQPLVILA--PHAVEALVKQPPDRWLSNARWTHYQALL-DYDRV 1128
Qy 601 STPLVRL-----AFNLVGDPIPGAE-TFYTDGSCNRQ 632
Db 1129 QFGPVVALNPATLLPLPEGLQHNLCDILAEAHGTRPDLTDQPLPADHTWYDGSLLQ 1188
Qy 633 SKGKAGYVDRGDKV--KKLEQ-TTNOQAELEAFAMAL-TDSGPKVNIIVDSQYVM-- 686
Db 1189 EGORKAGAAVTTETETIWAQKALPAGTSAQRAELIALTQALKMAEGKKNVYTDSDYAFAT 1248
Qy 687 -----GISASQPTSESK-IVNQIIEEMIKKEAIIYVAVWPAH-KIGNGNEVDH 733
Db 1249 AHITHGIYRRRGLLTSEGKEIKNKLDELALLKALFLPKRLSIITHCPGHQKHSAAERGNR 1308
Qy 734 LVSGQIRQ-----VLFLEKIEPQAEH-----EKY--- 758
Db 1309 MADQAARKAAITETPTDSTILLIENSSPYTSEHFTYVTDIKDLTKLCAIYDKTKKVVYQ 1368
Qy 759 -----HSNVKLSHKFGIP-----NLVAROIVNSCAOQ 788
Db 1369 GKVPMPDQFTFELDFELHQLTHLSFKMKALLERSHSPYMLNRDRTLKNITETCKACAQ 1428
Qy 789 -----KGEAIHQVNAELGT-WQMDCTHLEG-----KIIIVAVHVSAGFIEAIVP 833
Db 1429 VNASKSAVKGQGTVRGH---RPGTHWEIDFTEIKPGLYGYKLYLVLFIDTFSGWIEAFPTK 1485
Qy 834 QESGR-OTALFLKLKASRWPIH-LHTDNGANFTSQEVKVMVNWIGIEQSGVGPYPNQ 891
Db 1486 KETAKVVTKLLEIEIPRFGMPQVLTGDNCPAFVSKVSQVATVALLGLDWKLHCAYRQSS 1545
Qy 892 GVVYEAAMNHLKNOISRIEQANTIETIVLMAIHCMMNFKRRGIGDWTTPSERL----- 943
```

Search completed: September 25, 2003, 12:04:41
Job time : 33 secs

THIS PAGE BLANK (USPTO)

!!AA_MULTIPLE_ALIGNMENT 1.0
PileUp of: *.swissprot

Symbol comparison table: GenRunData:blosum62.cmp CompCheck: 1102

GapWeight: 8
GapLengthWeight: 2

sp.msrf MSF: 1165 Type: P September 25, 2003 10:17 Check: 5858

Name: pol_hv2d1	Len: 1165	Check: 9753	Weight: 1.00
Name: pol_hv2g1	Len: 1165	Check: 9239	Weight: 1.00
Name: pol_hv2be	Len: 1165	Check: 307	Weight: 1.00
Name: pol_hv2ca	Len: 1165	Check: 9370	Weight: 1.00
Name: pol_hv2ro	Len: 1165	Check: 4767	Weight: 1.00
Name: pol_hv2nz	Len: 1165	Check: 8741	Weight: 1.00
Name: pol_hv2kr	Len: 1165	Check: 8455	Weight: 1.00
Name: pol_hv2st	Len: 1165	Check: 9818	Weight: 1.00
Name: pol_hv2sb	Len: 1165	Check: 2098	Weight: 1.00
Name: pol_hv2s4	Len: 1165	Check: 6726	Weight: 1.00
Name: pol_hv2d2	Len: 1165	Check: 6584	Weight: 1.00

//

1	VLELWKGTL	GETVSTQKT	GLLEVVQVRT	HGKLPKQKTG	RFRDQPTGK	50
pol_hv2d1	-----	-----	-----	-----	-----	
pol_hv2g1	-----	-----	-----	-----	-----	
pol_hv2be	-----	-----	-----	-----	-----	
pol_hv2ca	-----	-----	-----	-----	-----	
pol_hv2ro	-----	-----	-----	-----	-----	
pol_hv2nz	-----	-----	-----	-----	-----	
pol_hv2kr	-----	-----	-----	-----	-----	
pol_hv2st	-----	-----	-----	-----	-----	
pol_hv2sb	-----	-----	-----	-----	-----	
pol_hv2s4	-----	-----	-----	-----	-----	
pol_hv2d2	-----	-----	-----	-----	-----	
51	AAPQLPRGSP	SSGADTNSPT	NRSSSG...	P	VGEIYAAREK	AERAEGETIQ
pol_hv2d1	-----	-----	-----	-----	-----	100
pol_hv2g1	-----	-----	-----	-----	-----	
pol_hv2be	-----	-----	-----	-----	-----	
pol_hv2ca	-----	-----	-----	-----	-----	
pol_hv2ro	-----	-----	-----	-----	-----	
pol_hv2nz	-----	-----	-----	-----	-----	
pol_hv2kr	-----	-----	-----	-----	-----	
pol_hv2st	-----	-----	-----	-----	-----	
pol_hv2sb	-----	-----	-----	-----	-----	
pol_hv2s4	-----	-----	-----	-----	-----	
pol_hv2d2	-----	-----	-----	-----	-----	
101	GGDGLTAPR	AGRDAPO	RG	DRGLATPOFS	LWKRPPVVTAF	IEDQPEVEVLL
pol_hv2d1	-----	-----	-----	-----	-----	150
pol_hv2g1	-----	-----	-----	-----	-----	
pol_hv2be	-----	-----	-----	-----	-----	
pol_hv2ca	-----	-----	-----	-----	-----	
pol_hv2ro	-----	-----	-----	-----	-----	
pol_hv2nz	-----	-----	-----	-----	-----	
pol_hv2kr	-----	-----	-----	-----	-----	
pol_hv2st	-----	-----	-----	-----	-----	
pol_hv2sb	-----	-----	-----	-----	-----	
pol_hv2s4	-----	-----	-----	-----	-----	
pol_hv2d2	-----	-----	-----	-----	-----	
151	GTGADDSIVA	GIELGDNVTP	KIVGGIGGFI	NTKEYKNVEI	KVLNKRVRAT	200
pol_hv2d1	-----	-----	-----	-----	-----	
pol_hv2g1	-----	-----	-----	-----	-----	
pol_hv2be	-----	-----	-----	-----	-----	
pol_hv2ca	-----	-----	-----	-----	-----	
pol_hv2ro	-----	-----	-----	-----	-----	
pol_hv2nz	-----	-----	-----	-----	-----	
pol_hv2kr	-----	-----	-----	-----	-----	
pol_hv2st	-----	-----	-----	-----	-----	
pol_hv2sb	-----	-----	-----	-----	-----	

pol_hv2Ro

90

-

IMTGTPTINI
FGRNII-LJTAGU
MSLNLPUAKN

pol_hv2s4	DTGADDSIVA	GIELGPNYTP	KIVGGIGGFI	NTKEYKDVKI	KVLGKVIKFT	
pol_hv2d2	DTGVDDSDIVA	GIELGNSYTP	KIVGGIGGFI	NTKEYKDVKI	EVVGRKVRAT	
201	IMTGDTPINI	FGNIIATLIG	MSLNLPUAKL	DPIKVTLPKG	KDGPRLKQWP	250
pol_hv2d1	IMTGDTPINI	FGNIIATLIG	MSLNLPUAKL	DPIKVTLPKG	KDGPRLKQWP	
pol_hv2g1	IMTGDTPINI	FGNIIATLIG	MSLNLPUAKI	EPIKVTLPKG	KDGPRLKQWP	
pol_hv2be	IMTGDTPINI	FGNIIATLIG	MSLNLPUAKI	EPIKVTLPKG	KDGPRLKQWP	
pol_hv2ca	IMTGDTPINI	FGNIIATLIG	MSLNLPUAKI	EPIKVTLPKG	KDGPRLKQWP	
pol_hv2ro	IMTGDTPINI	FGNIIATLIG	MSLNLPUAKI	EPIKVTLPKG	KDGPRLKQWP	
pol_hv2nz	IMTGDTPINI	FGNIIATLIG	MSLNLPUAKI	EPIKVTLPKG	KDGPRLKQWP	
pol_hv2kr	IMTGDTPINI	FGNIIATLIG	MSLNLPUAKI	EPIKVTLPKG	KDGPRLKQWP	
pol_hv2st	IMTGDTPINI	FGNIIATLIG	MSLNLPUAKI	EPIKVTLPKG	KDGPRLKQWP	
pol_hv2sb	IMTGDTPINI	FGNIIATLIG	MSLNLPUAKI	EPIKVTLPKG	KDGPRLKQWP	
pol_hv2s4	IMTGDTPINI	FGNIIATLIG	MSLNLPUAKI	EPIKVTLPKG	KDGPRLKQWP	
pol_hv2d2	IMTGDTPINI	FGNIIATLIG	MSLNLPUAKI	EPIKVTLPKG	KDGPRLKQWP	
251	LTRKEIEALK	EICEKMEREG	QLEEAPPTNP	YNTPTFAIKK	KDKNKKWRMLI	300
pol_hv2d1	LTRKEIEALK	EICEKMEREG	QLEEAPPTNP	YNTPTFAIKK	KDKNKKWRMLI	
pol_hv2g1	LTRKEIEALK	EICEKMEREG	QLEEAPPTNP	YNTPTFAIKK	KDKNKKWRMLI	
pol_hv2be	LTRKEIEALK	EICEKMEREG	QLEEAPPTNP	YNTPTFAIKK	KDKNKKWRMLI	
pol_hv2ca	LTRKEIEALK	EICEKMEREG	QLEEAPPTNP	YNTPTFAIKK	KDKNKKWRMLI	
pol_hv2ro	LTRKEIEALK	EICEKMEREG	QLEEAPPTNP	YNTPTFAIKK	KDKNKKWRMLI	
pol_hv2nz	LTRKEIEALK	EICEKMEREG	QLEEAPPTNP	YNTPTFAIKK	KDKNKKWRMLI	
pol_hv2kr	LTRKEIEALK	EICEKMEREG	QLEEAPPTNP	YNTPTFAIKK	KDKNKKWRMLI	
pol_hv2st	LTRKEIEALK	EICEKMEREG	QLEEAPPTNP	YNTPTFAIKK	KDKNKKWRMLI	
pol_hv2sb	LTRKEIEALK	EICEKMEREG	QLEEAPPTNP	YNTPTFAIKK	KDKNKKWRMLI	
pol_hv2s4	LSREKIIIALR	EICEKMEREG	QLEEAPPTNP	YNTPTFAIKK	KDKNKKWRMLI	
pol_hv2d2	LSREKILALK	EICEKMEREG	QLEEAPPTNP	YNTPTFAIKK	KDKNKKWRMLI	
301	DFRELNRVQT	DFTETIOLGIP	HPAGLAKKRR	ITVLVDGDAY	FSIPLHEDFR	350
pol_hv2d1	DFRELNRVQT	DFTETIOLGIP	HPAGLAKKRR	ITVLVDGDAY	FSIPLHEDFR	
pol_hv2g1	DFRELNRVQT	DFTETIOLGIP	HPAGLAKKRR	ITVLVDGDAY	FSIPLHEDFR	
pol_hv2be	DFRELNRVQT	DFTETIOLGIP	HPAGLAKKRR	ITVLVDGDAY	FSIPLHEDFR	
pol_hv2ca	DFRELNRVQT	DFTETIOLGIP	HPAGLAKKRR	ITVLVDGDAY	FSIPLHEDFR	
pol_hv2ro	DFRELNRVQT	DFTETIOLGIP	HPAGLAKKRR	ITVLVDGDAY	FSIPLHEDFR	
pol_hv2nz	DFRELNRVQT	DFTETIOLGIP	HPAGLAKKRR	ITVLVDGDAY	FSIPLHEDFR	
pol_hv2kr	DFRELNRVQT	DFTETIOLGIP	HPAGLAKKRR	ITVLVDGDAY	FSIPLHEDFR	
pol_hv2st	DFRELNRVQT	DFTETIOLGIP	HPAGLAKKRR	ITVLVDGDAY	FSIPLHEDFR	
pol_hv2sb	DFRELNRVQT	DFTETIOLGIP	HPAGLAKKRR	ITVLVDGDAY	FSIPLHEDFR	
pol_hv2s4	DFRELNRVQT	DFTETIOLGIP	HPAGLAKKRR	ITVLVDGDAY	FSIPLHEDFR	
pol_hv2d2	DFRELNRVQT	DFTETIOLGIP	HPAGLAKKRR	ITVLVDGDAY	FSIPLHEDFR	
351	QYTAFTLPSV	NNAEPGRKRV	YKVLPOGWKG	SPAIFQPMR	QILEPPERKAN	400
pol_hv2d1	QYTAFTLPSV	NNAEPGRKRV	YKVLPOGWKG	SPAIFQPMR	QILEPPERKAN	
pol_hv2g1	QYTAFTLPSV	NNAEPGRKRV	YKVLPOGWKG	SPAIFQPMR	QILEPPERKAN	
pol_hv2be	QYTAFTLPSV	NNAEPGRKRV	YKVLPOGWKG	SPAIFQPMR	QILEPPERKAN	
pol_hv2ca	QYTAFTLPSV	NNAEPGRKRV	YKVLPOGWKG	SPAIFQPMR	QILEPPERKAN	
pol_hv2ro	QYTAFTLPSV	NNAEPGRKRV	YKVLPOGWKG	SPAIFQPMR	QILEPPERKAN	
pol_hv2nz	QYTAFTLPSV	NNAEPGRKRV	YKVLPOGWKG	SPAIFQPMR	QILEPPERKAN	
pol_hv2kr	QYTAFTLPSV	NNAEPGRKRV	YKVLPOGWKG	SPAIFQPMR	QILEPPERKAN	
pol_hv2st	QYTAFTLPSV	NNAEPGRKRV	YKVLPOGWKG	SPAIFQPMR	QILEPPERKAN	
pol_hv2sb	QYTAFTLPSV	NNAEPGRKRV	YKVLPOGWKG	SPAIFQPMR	QILEPPERKAN	
pol_hv2s4	QYTAFTLPSV	NNAEPGRKRV	YKVLPOGWKG	SPAIFQPMR	QILEPPERKAN	
pol_hv2d2	QYTAFTLPSV	NNAEPGRKRV	YKVLPOGWKG	SPAIFQPMR	QILEPPERKAN	
401	PDVILIIQYMD	DILIASDRTG	LEHDKVVLQOL	KELLNGLGFS	TPDEKFQKDP	450
pol_hv2d1	PDVILIIQYMD	DILIASDRTG	LEHDKVVLQOL	KELLNGLGFS	TPDEKFQKDP	
pol_hv2g1	PDVILIIQYMD	DILIASDRTG	LEHDKVVLQOL	KELLNGLGFS	TPDEKFQKDP	
pol_hv2be	PDVILIIQYMD	DILIASDRTG	LEHDKVVLQOL	KELLNGLGFS	TPDEKFQKDP	
pol_hv2ca	PDVILIIQYMD	DILIASDRTG	LEHDKVVLQOL	KELLNGLGFS	TPDEKFQKDP	
pol_hv2ro	PDVILIIQYMD	DILIASDRTG	LEHDKVVLQOL	KELLNGLGFS	TPDEKFQKDP	
pol_hv2nz	PDVILIIQYMD	DILIASDRTG	LEHDKVVLQOL	KELLNGLGFS	TPDEKFQKDP	
pol_hv2kr	PDVILIIQYMD	DILIASDRTG	LEHDKVVLQOL	KELLNGLGFS	TPDEKFQKDP	
pol_hv2st	PDVILIIQYMD	DILIASDRTG	LEHDKVVLQOL	KELLNGLGFS	TPDEKFQKDP	
pol_hv2sb	PDVILIIQYMD	DILIASDRTG	LEHDKVVLQOL	KELLNGLGFS	TPDEKFQKDP	
pol_hv2s4	PDVILIIQYMD	DILIASDRTG	LEHDKVVLQOL	KELLNGLGFS	TPDEKFQKDP	
pol_hv2d2	PDVILIIQYMD	DILIASDRTG	LEHDKVVLQOL	KELLNGLGFS	TPDEKFQKDP	
451	PFOWMGYELW	PTKWKLOKIQ	LPOKEIWTVN	DIQKLVGVLN	WAAQIYPGIK	500
pol_hv2d1	PFOWMGYELW	PTKWKLOKIQ	LPOKEIWTVN	DIQKLVGVLN	WAAQIYPGIK	
pol_hv2g1	PFOWMGYELW	PTKWKLOKIQ	LPOKEIWTVN	DIQKLVGVLN	WAAQIYPGIK	
pol_hv2be	PFOWMGYELW	PTKWKLOKIQ	LPOKEIWTVN	DIQKLVGVLN	WAAQIYPGIK	
pol_hv2ca	PFOWMGYELW	PTKWKLOKIQ	LPOKEIWTVN	DIQKLVGVLN	WAAQIYPGIK	

pol_hv2ro PYHMGYELW PTKWKLOKIQ LPOKEIWTNVN DIQKLVGVNL WAAOLYPGIK
pol_hv2nz PYHMGYELW PTKWKLOKIQ LPOKEVWTVN DIQKLVGVNL WAAQIYPGIK
pol_hv2st PYHMGYELW PTKWKLOKIQ LPOKEVWTVN DIQKLVGVNL WAAQIYPGIK
pol_hv2sb PYHMGYELW PTKWKLOKIQ LPOKEVWTVN DIQKLVGVNL WAAQIYPGIK
pol_sivs4 PYHMGYELW PTKWKLOKIQ LPOKEVWTVN DIQKLVGVNL WAAQIYPGIK
pol_hv2d2 PYHMGYELW PTKWKLOKIQ LPEKEVWTVN AIOKLVGVNL WAAOLFTPGIK

501
pol_hv2d1 TKHLCLRLNG KMTLITEEVQW TELAEAELEE NKIILSQBOE GSYQOESEEL
pol_hv2g1 TKHLCLRLNG KMTLITEEVQW TELAEAELEE NKIILSQBOE GYTYOESEEL
pol_hv2be TKHLCLRLNG KMTLITEEVQW TELAEAELEE NKIILSQBOE GYTYOESEEL
pol_hv2ca TKHLCLRLNG KMTLITEEVQW TELAEAELEE NKIILSQBOE GYTYOESEEL
pol_hv2n2 TKHLCLRLNG KMTLITEEVQW TELAEAELEE NKIILSQBOE GYTYOESEEL
pol_hv2n2 TKHLCLRLNG KMTLITEEVQW TELAEAELEE NKIILSQBOE GYTYOESEEL
pol_hv2st TKHLCLRLNG KMTLITEEVQW TELAEAELEE NKIILSQBOE GYTYOESEEL
pol_hv2sb TKHLCLRLNG KMTLITEEVQW TELAEAELEE NKIILSQBOE GYTYOESEEL
pol_sivs4 TKHLCLRLNG KMTLITEEVQW TELAEAELEE NKIILSQBOE GYTYOESEEL
pol_hv2d2 TKHLCLRLNG KMTLITEEVQW TELAEAELEE NKIILSQBOE GYTYOESEEL

551
pol_hv2d1 EATVKSODN QWATYKHQGE RVLKVGKGYAK IKNTHTNGVR LLAQVVOKIG
pol_hv2g1 EATVKSODN QWATYKHQGE RVLKVGKGYAK IKNTHTNGVR LLAQVVOKIG
pol_hv2be EATVKSODN QWATYKHQGE RVLKVGKGYAK IKNTHTNGVR LLAQVVOKIG
pol_hv2ca EATVKSODN QWATYKHQGE RVLKVGKGYAK IKNTHTNGVR LLAQVVOKIG
pol_hv2n2 EATVKSODN QWATYKHQGE RVLKVGKGYAK IKNTHTNGVR LLAQVVOKIG
pol_hv2n2 EATVKSODN QWATYKHQGE RVLKVGKGYAK IKNTHTNGVR LLAQVVOKIG
pol_hv2st EATVKSODN QWATYKHQGE RVLKVGKGYAK IKNTHTNGVR LLAQVVOKIG
pol_hv2sb EATVKSODN QWATYKHQGE RVLKVGKGYAK IKNTHTNGVR LLAQVVOKIG
pol_sivs4 EATVKSODN QWATYKHQGE RVLKVGKGYAK IKNTHTNGVR LLAQVVOKIG
pol_hv2d2 EATVKSODN QWATYKHQGE RVLKVGKGYAK IKNTHTNGVR LLAQVVOKIG

601
pol_hv2d1 KEALVINGRV PKFHLPVPERE TWQWMDNDYV QVTWVPEWDF VSTPPLVRLT
pol_hv2g1 KEALVINGRV PKFHLPVPERE TWQWMDNDYV QVTWVPEWDF VSTPPLVRLT
pol_hv2be KEALVINGRV PKFHLPVPERE TWQWMDNDYV QVTWVPEWDF VSTPPLVRLT
pol_hv2ca KEALVINGRV PKFHLPVPERE TWQWMDNDYV QVTWVPEWDF VSTPPLVRLT
pol_hv2n2 KEALVINGRV PKFHLPVPERE TWQWMDNDYV QVTWVPEWDF VSTPPLVRLT
pol_hv2n2 KEALVINGRV PKFHLPVPERE TWQWMDNDYV QVTWVPEWDF VSTPPLVRLT
pol_hv2st KEALVINGRV PKFHLPVPERE TWQWMDNDYV QVTWVPEWDF VSTPPLVRLT
pol_hv2sb KEALVINGRV PKFHLPVPERE TWQWMDNDYV QVTWVPEWDF VSTPPLVRLT
pol_sivs4 KEALVINGRV PKFHLPVPERE TWQWMDNDYV QVTWVPEWDF VSTPPLVRLT
pol_hv2d2 KEALVINGRV PKFHLPVPERE TWQWMDNDYV QVTWVPEWDF VSTPPLVRLT

650
pol_hv2d1 FNLVGDPIPG TETFTYTDGSC NROSKEGKAG YVTDGRDRV RVLEQTTNQQ
pol_hv2g1 FNLVGDPIPG TETFTYTDGSC NROSKEGKAG YVTDGRDRV RVLEQTTNQQ
pol_hv2be FNLVGDPIPG TETFTYTDGSC NROSKEGKAG YVTDGRDRV RVLEQTTNQQ
pol_hv2ca FNLVGDPIPG TETFTYTDGSC NROSKEGKAG YVTDGRDRV RVLEQTTNQQ
pol_hv2n2 FNLVGDPIPG TETFTYTDGSC NROSKEGKAG YVTDGRDRV RVLEQTTNQQ
pol_hv2n2 FNLVGDPIPG TETFTYTDGSC NROSKEGKAG YVTDGRDRV RVLEQTTNQQ
pol_hv2st FNLVGDPIPG TETFTYTDGSC NROSKEGKAG YVTDGRDRV RVLEQTTNQQ
pol_hv2sb FNLVGDPIPG TETFTYTDGSC NROSKEGKAG YVTDGRDRV RVLEQTTNQQ
pol_sivs4 FNLVGDPIPG TETFTYTDGSC NROSKEGKAG YVTDGRDRV RVLEQTTNQQ
pol_hv2d2 FNLVGDPIPG TETFTYTDGSC NROSKEGKAG YVTDGRDRV RVLEQTTNQQ

700
pol_hv2d1 FNLVGDPIPG TETFTYTDGSC NROSKEGKAG YVTDGRDRV RVLEQTTNQQ
pol_hv2g1 FNLVGDPIPG TETFTYTDGSC NROSKEGKAG YVTDGRDRV RVLEQTTNQQ
pol_hv2be FNLVGDPIPG TETFTYTDGSC NROSKEGKAG YVTDGRDRV RVLEQTTNQQ
pol_hv2ca FNLVGDPIPG TETFTYTDGSC NROSKEGKAG YVTDGRDRV RVLEQTTNQQ
pol_hv2n2 FNLVGDPIPG TETFTYTDGSC NROSKEGKAG YVTDGRDRV RVLEQTTNQQ
pol_hv2n2 FNLVGDPIPG TETFTYTDGSC NROSKEGKAG YVTDGRDRV RVLEQTTNQQ
pol_hv2st FNLVGDPIPG TETFTYTDGSC NROSKEGKAG YVTDGRDRV RVLEQTTNQQ
pol_hv2sb FNLVGDPIPG TETFTYTDGSC NROSKEGKAG YVTDGRDRV RVLEQTTNQQ
pol_sivs4 FNLVGDPIPG TETFTYTDGSC NROSKEGKAG YVTDGRDRV RVLEQTTNQQ
pol_hv2d2 FNLVGDPIPG TETFTYTDGSC NROSKEGKAG YVTDGRDRV RVLEQTTNQQ

750
pol_hv2d1 AELEAFAMAL ADSGPKVNI I VDSQYVNGIV AGQPTESESR INVQIIEEMI
pol_hv2g1 AELEAFAMAL ADSGPKVNI I VDSQYVNGIV AGQPTESESR INVQIIEEMI
pol_hv2be AELEAFAMAL ADSGPKVNI I VDSQYVNGIV AGQPTESESR INVQIIEEMI
pol_hv2ca AELEAFAMAL ADSGPKVNI I VDSQYVNGIV AGQPTESESR INVQIIEEMI
pol_hv2n2 AELEAFAMAL ADSGPKVNI I VDSQYVNGIV AGQPTESESR INVQIIEEMI
pol_hv2n2 AELEAFAMAL ADSGPKVNI I VDSQYVNGIV AGQPTESESR INVQIIEEMI
pol_hv2st AELEAFAMAL ADSGPKVNI I VDSQYVNGIV AGQPTESESR INVQIIEEMI
pol_hv2sb AELEAFAMAL ADSGPKVNI I VDSQYVNGIV AGQPTESESR INVQIIEEMI
pol_sivs4 AELEAFAMAL ADSGPKVNI I VDSQYVNGIV AGQPTESESR INVQIIEEMI
pol_hv2d2 AELEAFAMAL ADSGPKVNI I VDSQYVNGIV AGQPTESESR INVQIIEEMI

751
pol_hv2d1 KKEAVYVAVV PAHKGIGGNQ EVDHLVSQGI ROVLFEKIE PAOEHEKHYH 800
pol_hv2g1 KKEAVYVAVV PAHKGIGGNQ EVDHLVSQGI ROVLFEKIE PAOEHEKHYH
pol_hv2be KKEAVYVAVV PAHKGIGGNQ EVDHLVSQGI ROVLFEKIE PAOEHEKHYH
pol_hv2ca KKEAVYVAVV PAHKGIGGNQ EVDHLVSQGI ROVLFEKIE PAOEHEKHYH
pol_hv2n2 KKEAVYVAVV PAHKGIGGNQ EVDHLVSQGI ROVLFEKIE PAOEHEKHYH
pol_hv2n2 KKEAVYVAVV PAHKGIGGNQ EVDHLVSQGI ROVLFEKIE PAOEHEKHYH
pol_hv2st KKEAVYVAVV PAHKGIGGNQ EVDHLVSQGI ROVLFEKIE PAOEHEKHYH
pol_hv2sb KKEAVYVAVV PAHKGIGGNQ EVDHLVSQGI ROVLFEKIE PAOEHEKHYH
pol_sivs4 KKEAVYVAVV PAHKGIGGNQ EVDHLVSQGI ROVLFEKIE PAOEHEKHYH
pol_hv2d2 KKEAVYVAVV PAHKGIGGNQ EVDHLVSQGI ROVLFEKIE PAOEHEKHYH

801
pol_hv2d1 SNIKELTHKF GIPOLVAROI VNTCAOQOK GEAIHGOVNA EIGVWQMDCT 850
pol_hv2g1 SNIKELTHKF GIPOLVAROI VNTCAOQOK GEAIHGOVNA EIGVWQMDCT
pol_hv2be SNIKELTHKF GIPOLVAROI VNTCAOQOK GEAIHGOVNA EIGVWQMDCT
pol_hv2ca SNIKELTHKF GIPOLVAROI VNTCAOQOK GEAIHGOVNA EIGVWQMDCT
pol_hv2n2 SNIKELTHKF GIPOLVAROI VNTCAOQOK GEAIHGOVNA EIGVWQMDCT
pol_hv2n2 SNIKELTHKF GIPOLVAROI VNTCAOQOK GEAIHGOVNA EIGVWQMDCT
pol_hv2st SNIKELTHKF GIPOLVAROI VNTCAOQOK GEAIHGOVNA EIGVWQMDCT
pol_hv2sb SNIKELTHKF GIPOLVAROI VNTCAOQOK GEAIHGOVNA EIGVWQMDCT
pol_sivs4 SNIKELTHKF GIPOLVAROI VNTCAOQOK GEAIHGOVNA EIGVWQMDCT
pol_hv2d2 SNIKELTHKF GIPOLVAROI VNTCAOQOK GEAIHGOVNA EIGVWQMDCT

851
pol_hv2d1 HLEGKIIIVA VHVASGFIEA EVIPOSGRQ TALFLKLKAS RWPITHLHTD 900
pol_hv2g1 HLEGKIIIVA VHVASGFIEA EVIPOSGRQ TALFLKLKAS RWPITHLHTD
pol_hv2be HLEGKIIIVA VHVASGFIEA EVIPOSGRQ TALFLKLKAS RWPITHLHTD
pol_hv2ca HLEGKIIIVA VHVASGFIEA EVIPOSGRQ TALFLKLKAS RWPITHLHTD
pol_hv2n2 HLEGKIIIVA VHVASGFIEA EVIPOSGRQ TALFLKLKAS RWPITHLHTD
pol_hv2n2 HLEGKIIIVA VHVASGFIEA EVIPOSGRQ TALFLKLKAS RWPITHLHTD
pol_hv2st HLEGKIIIVA VHVASGFIEA EVIPOSGRQ TALFLKLKAS RWPITHLHTD
pol_hv2sb HLEGKIIIVA VHVASGFIEA EVIPOSGRQ TALFLKLKAS RWPITHLHTD
pol_sivs4 HLEGKIIIVA VHVASGFIEA EVIPOSGRQ TALFLKLKAS RWPITHLHTD
pol_hv2d2 HLEGKIIIVA VHVASGFIEA EVIPOSGRQ TALFLKLKAS RWPITHLHTD

901
pol_hv2d1 NGSNFTSOEV KMAVWVGIE QSFQVVPYNPQ SOGVVEAMNH HLKNOISRIR 950
pol_hv2g1 NGSNFTSOEV KMAVWVGIE QSFQVVPYNPQ SOGVVEAMNH HLKNOISRIR
pol_hv2be NGSNFTSOEV KMAVWVGIE QSFQVVPYNPQ SOGVVEAMNH HLKNOISRIR
pol_hv2ca NGSNFTSOEV KMAVWVGIE QSFQVVPYNPQ SOGVVEAMNH HLKNOISRIR
pol_hv2n2 NGSNFTSOEV KMAVWVGIE QSFQVVPYNPQ SOGVVEAMNH HLKNOISRIR
pol_hv2n2 NGSNFTSOEV KMAVWVGIE QSFQVVPYNPQ SOGVVEAMNH HLKNOISRIR
pol_hv2st NGSNFTSOEV KMAVWVGIE QSFQVVPYNPQ SOGVVEAMNH HLKNOISRIR
pol_hv2sb NGSNFTSOEV KMAVWVGIE QSFQVVPYNPQ SOGVVEAMNH HLKNOISRIR
pol_sivs4 NGSNFTSOEV KMAVWVGIE QSFQVVPYNPQ SOGVVEAMNH HLKNOISRIR
pol_hv2d2 NGSNFTSOEV KMAVWVGIE QSFQVVPYNPQ SOGVVEAMNH HLKNOISRIR

951
pol_hv2d1 EQANTTIETIV LMAVHCNMFK RRGIGDMTP AERLINMITT EOEIQFLQK 1000
pol_hv2g1 EQANTTIETIV LMAVHCNMFK RRGIGDMTP AERLINMITT EOEIQFLQK
pol_hv2be EQANTTIETIV LMAVHCNMFK RRGIGDMTP AERLINMITT EOEIQFLQK
pol_hv2ca EQANTTIETIV LMAVHCNMFK RRGIGDMTP AERLINMITT EOEIQFLQK
pol_hv2n2 EQANTTIETIV LMAVHCNMFK RRGIGDMTP AERLINMITT EOEIQFLQK
pol_hv2n2 EQANTTIETIV LMAVHCNMFK RRGIGDMTP AERLINMITT EOEIQFLQK
pol_hv2st EQANTTIETIV LMAVHCNMFK RRGIGDMTP AERLINMITT EOEIQFLQK
pol_hv2sb EQANTTIETIV LMAVHCNMFK RRGIGDMTP AERLINMITT EOEIQFLQK
pol_sivs4 EQANTTIETIV LMAVHCNMFK RRGIGDMTP AERLINMITT EOEIQFLQK
pol_hv2d2 EQANTTIETIV LMAVHCNMFK RRGIGDMTP AERLINMITT EOEIQFLQK

1001
pol_hv2d1 NSNFKKFOVY YRGRDOLWK GPCELLWKGD GAVIVKVGAD IKVPRRKAK 1050
pol_hv2g1 NSNFKKFOVY YRGRDOLWK GPCELLWKGD GAVIVKVGAD IKVPRRKAK
pol_hv2be NSNFKKFOVY YRGRDOLWK GPCELLWKGD GAVIVKVGAD IKVPRRKAK
pol_hv2ca NSNFKKFOVY YRGRDOLWK GPCELLWKGD GAVIVKVGAD IKVPRRKAK
pol_hv2n2 NSNFKKFOVY YRGRDOLWK GPCELLWKGD GAVIVKVGAD IKVPRRKAK
pol_hv2n2 NSNFKKFOVY YRGRDOLWK GPCELLWKGD GAVIVKVGAD IKVPRRKAK
pol_hv2st NSNFKKFOVY YRGRDOLWK GPCELLWKGD GAVIVKVGAD IKVPRRKAK
pol_hv2sb NSNFKKFOVY YRGRDOLWK GPCELLWKGD GAVIVKVGAD IKVPRRKAK
pol_sivs4 NSNFKKFOVY YRGRDOLWK GPCELLWKGD GAVIVKVGAD IKVPRRKAK
pol_hv2d2 NSNFKKFOVY YRGRDOLWK GPCELLWKGD GAVIVKVGAD IKVPRRKAK

```
pol_hv2st  NSKLQNRVY  FREGRDQLWK  GPGELLWKGD  GAVIVKVGAD  IKIIPRRKAK
pol_hv2sb  NSKLQNRVY  FREGRDQLWK  GPGELLWKGD  GAVIVKVGTD  IKVIPRRKAK
pol_si4    NSKFKNRVY  YREGRDQLWK  GPGELLWKGE  GAVILKVGTE  IKVVPRRKAK
pol_hv2d2  NLKFNQFVY  YREGRDQLWK  GPGELLWKGE  GAVILKVGTE  IKVVPRRKAK

1051
pol_hv2d1  IIRDYGRQE  LDSSHLEGA  R.EDGEVA--  -----  1100
pol_hv2g1  IIRDYGRQE  LD.  SSHLEGA  REEDGEVA--  -----
pol_hv2be  IIRDYGRQE  LDSSHLEGA  R.EDGEVACP  CQVPEIONKR  PRGGALCSPP
pol_hv2ca  IIRDYGRQE  LDSSHLEGA  R.EDGEVA--  -----
pol_hv2ro  IIRDYGRQE  MDSSGHLEGA  R.EDGEVA--  -----
pol_hv2nz  IIRDYGRQE  MDSSGHLEGA  R.EDGEVA--  -----
pol_hv2kr  IIRDYGRQE  VSSSHLEGT  R.EDGEVA--  -----
pol_hv2st  IIRDYGRQE  MDSSGNLEGA  R.EDGEVA--  -----
pol_hv2sb  IIRDYGRQE  MDSSGNLEGA  R.EDGEVA--  -----
pol_si4s4  IIRDYGGGKE  LDSSGHLEDT  .GEAREVA--  -----
pol_hv2d2  IIRHYGGGKG  LDSSADMET  R.QAREMAQS  D-----

1101
pol_hv2d1  -----  -----  -----  -----  1150
pol_hv2g1  -----  -----  -----  -----
pol_hv2be  QGGMGWVDLQ  OGNIPTRKK  SSRNTGILEP  NTRKRMALLS  CSKINLVYRK
pol_hv2ca  -----  -----  -----  -----
pol_hv2ro  -----  -----  -----  -----
pol_hv2nz  -----  -----  -----  -----
pol_hv2kr  -----  -----  -----  -----
pol_hv2st  -----  -----  -----  -----
pol_hv2sb  -----  -----  -----  -----
pol_si4s4  -----  -----  -----  -----
pol_hv2d2  -----  -----  -----  -----

1151
pol_hv2d1  -----  1165
pol_hv2g1  -----
pol_hv2be  VLDRCYPRLC  RHPNT
pol_hv2ca  -----
pol_hv2ro  -----
pol_hv2nz  -----
pol_hv2kr  -----
pol_hv2st  -----
pol_hv2sb  -----
pol_si4s4  -----
pol_hv2d2  -----
```

THIS PAGE BLANK (USPTO)

!!AA_MULTIPLE_ALIGNMENT 1.0
Pileup of: *.sp_rvirus

Symbol comparison table: GenRunData:blosum62.cmp CompCheck: 1102

GapWeight: 8
GapLengthWeight: 2

pileup.msf MSF: 1645 Type: P September 25, 2003 10:13 Check: 2437 ..

Name: q76053 Len: 1645 Check: 4352 Weight: 1.00
Name: q90066 Len: 1645 Check: 2138 Weight: 1.00
Name: q77162 Len: 1645 Check: 4711 Weight: 1.00
Name: q77249 Len: 1645 Check: 3717 Weight: 1.00
Name: q73194 Len: 1645 Check: 7255 Weight: 1.00
Name: q78654 Len: 1645 Check: 3402 Weight: 1.00
Name: q76630 Len: 1645 Check: 4584 Weight: 1.00
Name: q9ytuo Len: 1645 Check: 9112 Weight: 1.00
Name: q78655 Len: 1645 Check: 3199 Weight: 1.00
Name: q89928 Len: 1645 Check: 4930 Weight: 1.00
Name: q9ib19 Len: 1645 Check: 9116 Weight: 1.00
Name: q9dsk7 Len: 1645 Check: 5921 Weight: 1.00

//

1
q76053 50
q90066
q77162
q77249
q73194
q78654
q76630
q9ytuo
q78655
q89928
q9ib19
q9dsk7
TSTCYMRHCN TSVIKESCDK HYWDAMKPRY CAPPGFALLR CNDNTYSGFE

51
q76053 100
q90066
q77162
q77249
q73194
q78654
q76630
q9ytuo
q78655
q89928
q9ib19
q9dsk7
PNC SKVVATT CTRMMETQTS TWFGFNGSRA ENRTYIYWHS KDNRTIISLN

101
q76053 150
q90066
q77162
q77249
q73194
q78654
q76630
q9ytuo
q78655
q89928
q9ib19
q9dsk7
SFYNLTWHCK RPKGNTVVP I TLM SGLVFHS QPINKRPRQA KCWFGGNWTG

151
q76053 200
q90066
q77162
q77249
q73194

q78654
q76630
q9ytuo
q78655
q89928
q9ib19
q9dsk7
AMQEVKOTLA KTSQVORNHD TKKINFARPG MGSDEPVVYM WTNCRGEFLY

201
q76053 250
q90066
q77162
q77249
q73194
q78654
q76630
q9ytuo
q78655
q89928
q9ib19
q9dsk7
CNMTWFLNWV ENRTNTTOHN YAPCHIKOII NAWHKVGKYY YLPPREGELT

251
q76053 300
q90066
q77162
q77249
q73194
q78654
q76630
q9ytuo
q78655
q89928
q9ib19
q9dsk7
CNSTVTSIIA NIDTGANQIN ITFSAEVSDL YRLELDYKL IEITPIGFAP

301
q76053 350
q90066
q77162
q77249
q73194
q78654
q76630
q9ytuo
q78655
q89928
q9ib19
q9dsk7
TNVKRYSSTP VRNKRGVFVL GFLGFLATAG AAMGTASLTL SAQSRTLLAG

351
q76053 400
q90066
q77162
q77249
q73194
q78654
q76630
q9ytuo
q78655
q89928
q9ib19
q9dsk7
IVQOQOQLLD VVKRQOELLR LTVWGTKNLQ ARVTAIEKYL KHAQALNSMG

401
q76053 450
q90066
q77162
q77249
q73194
q78654
q76630
q9ytuo

q78655	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
--------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

q9dsk7	701	750			
q76053	WTVAYIEGQP	DSIVAGIELG	NNYSPKIVGG	IGGFINTKEY	750
q90066	WTVAYIEGQP	VEVLDDTGAD	DSIVAGIELG	NNYSPKIVGG	IGGFINTKEY
q77162	WTVAYIEGQP	VEVLDDTGAD	DSIVAGIELG	DNTPKIVGG	IGGFINTKEY
q77249	WTVAYIEGQP	VEVLDDTGAD	DSIVAGIELG	DNTPKIVGG	IGGFINTKEY
q73194	WTVAHIEGQP	VEVLDDTGAD	DSIVAGIELG	SNYSPKIVGG	IGGFINTKEY
q78654	WTVAHIEGQP	VEVLDDTGAD	DSIVAGIELG	CXYSPKIVGG	IGGFINTKEY
q76630	WTVAHIEGQP	VEVLDDTGAD	DSIVAGIELG	SNYSPKIVGG	IGGFINTKEY
q9YU00	WTVAYIEGQP	VEVLDDTGAD	DSIVAGIELG	SNYTPKIVGG	IGGFINTKEY
q78655	WTVAYIEGQP	VEVLDDTGAD	DSIVAGIELG	SNYTPKIVGG	IGGFINTKEY
q89928	WVKATIEGQS	VEVLDDTGAD	DSIVAGIELG	SNYTPKIVGG	IGGFINTKEY
q91b19	WTVAYIESQP	VEVLDDTGAD	DSIVAGIELG	SNYSPKIVGG	IGGFINTKEY
q9dsk7					
751	800				
q76053	KNVEIEVLNK	RVRAITMTGD	TPNIFGRNI	L7ALGMSLNL	800
q90066	KNVEIEVLNK	KVRATIMTGD	TPNIFGRNI	L7ALGMSLNL	
q77162	KNVEIKVLNK	RVRAITMTGD	TPNIFGRNI	L7ALGMSLNL	
q77249	KNVEIEVLNK	RVRAITMTGD	TPNIFGRNI	L7ALGMSLNL	
q73194	KNVEIEVLNK	RVRAITMTGD	TPNIFGRNI	L7ALGMSLNL	
q78654	KDVEIEVLGR	RVRAITMTGD	TPNIFGRNV	L7ALGMSLNL	PVAKTEPIKI
q76630	KNVEIKVLNK	RVRAITMTGD	TPNIFGRNI	L7ALGMSLNL	PVAKTEPIKI
q9YU00	EDVEIKVLNK	RVRAITMTGD	TPNIFGRNI	L7ALGMSLNL	PVAKTEPIEV
q78655	KNVKEIELLNK	RVRAITMTGD	TPNIFGRNI	LNSLGMTLNF	PVARTPEVKV
q89928	KNVEIEVGK	RVRAITMTGD	TPNIFGRNI	L7ALGMSLNF	PLAKVEPVKV
q91b19	KNVEIKVGK	RQAIVMTGD	TPNIFGRNI		
q9dsk7					
801	850				
q76053					
q90066					
q77162					
q77249					
q73194	MUKPGKDGR	LQWPLTKK	IEALKEICEK	MEKEGLEEA	PPTNPNYPTPT
q78654	MUKPGKDGR	LQWPLTKK	IEALKEICEK	MEKEGLEEA	PPTNPNYPTPT
q76630	MUKPGKDGR	LQWPLTKK	IEALKEICEK	MEKEGLEEA	PPTNPNYPTPT
q9YU00	MUKPGKDGR	LQWPLTKK	IEALKEICEK	MEKEGLEEA	PPTNPNYPTPT
q78655					
q89928	OLKPEKDGR	LQWPLSKK	ILALKEICEK	MEKEGLEEA	PPTNPNYNSPT
q91b19	RUKPGMDGR	LQWPLSKK	IQALREICDK	MEKEGLEEA	PPTNPNYPTPT
q9dsk7					
851	900				
q76053					
q90066					
q77162					
q77249					
q73194	FAIRKKDKNK	WRMLIDFREL	NKVTDQFTEI	OLGIPHPAGL	AKKRITVILD
q78654	FAIRKKDKNK	WRMLIDFREL	NKVTDQFTEI	OLGIPHPAGL	AKKRITVILD
q76630	FAIRKKDKNK	WRMLIDFREL	NKVTDQFTEI	OLGIPHPAGL	AKKRITVILD
q9YU00	FAIRKKDKNK	WRMLIDFREL	NKVTDQFTEI	OLGIPHPAGL	AKKRITVILD
q78655					
q89928	FAIRKKDKNK	WRMLIDFREL	NKVTDQFTEV	OLGIPHPAGL	AKKRITVILD
q91b19	FAIRKKDKNK	WRMLIDFREL	NKVTDQFTEV	OLGIPHPAGL	AKKRITVILD
q9dsk7					
901	950				
q76053					
q90066					
q77162					
q77249					
q73194	VGDAFYSIPL	HEDFRQYTF	TLPSVNNAEP	GKRYIKVLP	QWKGSPAIF
q78654	VGDAFYSIPL	HEDFRQYTF	TLPSVNNAEP	GKRYIKVLP	QWKGSPAIF
q76630	VGDAFYSIPL	HEDFRQYTF	TLPSVNNAEP	GKRYIKVLP	QWKGSPAIF
q9YU00	VGDAFYSIPL	HEDFRQYTF	TLPSVNNAEP	GKRYIKVLP	QWKGSPAIF
q78655					
q89928	VGDAFYSIPL	DPDEFQYTF	TLPSVNNAEP	GKRYIKVLP	QWKGSPAIF
q91b19	IGDAFYSIPL	DVDFQYTF	TLPSVNNAEP	GKRYIKVLP	QWKGSPAIF
q9dsk7					
951	1000				

```

q76053 -----
q90066 -----
q77162 -----
q77249 -----
q73194 QYTMROVLEP FRKANSVII IOYMDILIA SDRDLEHDK VYLQKELLN
q78654 -----
q76630 QYTMROVLEP FRKANPDVIL IOYMDILIA SDRDLEHDK VYLQKELLN
q9ytuo QYTMROVLEP FRKANODVIL IOYMDILIA SDRDLEHDK VYLQKELLN
q78655 -----
q89928 QYTMKVLEP FRKANNDVTI IOYMDILIA SDRDLEHDK VYSLKELLN
q91b19 QYTMKVLEP FRKANEDXII IOYMDILIA SDRDLEHDK VYLQKELLN
q9dsk7 -----

1001 -----
q76053 -----
q90066 -----
q77162 -----
q77249 -----
q73194 NLGFSPPDEK FOKDPPYRWM GYELWPTKWK LOKIQLPOKE VWTVNDIOKL
q78654 -----
q76630 GLGFSPPDEK FOKDPPYQWM GYELWPTKWK LOKIQLPOKE VWTVNDIOKL
q9ytuo GLGFSPPDEK FOKDPPYKWM GYELWPTKWK LOKIQLPOKE VWTVNDIOKL
q78655 -----
q89928 NMGFSPEEK FOKDPPYKWM GYELWPKKWK LOKIQLPEKE VWTVNDIOKL
q91b19 KYGFSPPDEK FOKDPPYQWM GYELWPKKWK LOKINLPEKE VWTVNDIOKL
q9dsk7 -----

1051 -----
q76053 -----
q90066 -----
q77162 -----
q77249 -----
q73194 VGVLNWAAOI YPGIKTKHLC RLIRGRMTLT EEVQWTELA EBLEENRIIL
q78654 -----
q76630 VGVLNWAAOI YPGIKTRNLC RLIRGRMTLT EEVQWTELA EBLEENKIIL
q9ytuo VGVLNWAAOI YPGIKTKHLC RLIRGRMTLT EGQWTELA EBLEENRIIL
q78655 -----
q89928 VGVLNWAAOI YPGIKTRHLC KLIRGRMTLT EEVQWTELA EBLEENKIIL
q91b19 VGVLNWAAOI YPGIKTRNLC KLIRGRMTLT EEVQWTELA EBLEENKIIL
q9dsk7 -----

1101 -----
q76053 -----
q90066 -----
q77162 -----
q77249 -----
q73194 SOEQEGHYO EKELEATVQ KDQDNQWYK IHQEGKILV GYAKIKHTH
q78654 -----
q76630 SOEQEGCYO EKELEATVQ KDQDNQWYK IHQGGKILV GYAKVKNTH
q9ytuo SOEQEGHYO EKELEATVQ KDQDNQWYK IHQEGKILV EYAKMKHNT
q78655 -----
q89928 EOEQEGSYK EGVPLEATVQ KNLANQWYK IHQGGKILV GYAKVKNTH
q91b19 XOEQEGAYK EKEPLEATVQ KNLDNQWYK IHQGGKILV GYAKIKHTH
q9dsk7 -----

1151 -----
q76053 -----
q90066 -----
q77162 -----
q77249 -----
q73194 TNGVKLLAOV VOKIGKEALV IXGRIPFHL PVEREWQWQ WDNVQWQWVI
q78654 -----
q76630 TNGVRLAOV VOKIGKEALV IWGRIPFHL PVERDTEWQ WDNVQWQWVI
q9ytuo TNGVRLAOV VOKIGKEALV IWGRIPFHL PVERETWQW WDDYQWQWVI
q78655 -----
q89928 TNGVRLAHV VOKIGKEALV IWGEIPFHL PVERETWQW WTDYQWQWVI
q91b19 TNGVRLAHV VOKIGKEALV IWKLPPFHL PVERDTEWQ WTDYQWQWVI
q9dsk7 -----

1201 -----
q76053 -----
q90066 -----
q77162 -----
q77249 -----
q73194 -----
q78654 -----
q76630 -----
q9ytuo -----
q89928 -----
q91b19 -----
q9dsk7 -----

1250 -----
q76053 -----
q90066 -----
q77162 -----
q77249 -----
q73194 -----
q78654 -----

```

```

q77249 -----
q73194 PWDFVSTPP LVRLAFNLVG DPILPGTFFY TDGSCNRQSK EGKAGYVTD
q78654 -----
q76630 PWDFVSTPP LVRLAFNLV DPILGAEFFY TDGSCNKQSR EGKAGYVTD
q9ytuo PWDFVSTPP LVRLAFNLV DPILGAEFFY TDGPCNRQSK EGKAGYVTD
q78655 -----
q89928 PWDFVSTPP LVRLAFNLV DPILGAEFFY TDGSCNKQSK EGKAGYVTD
q91b19 PWDFVSTPP LVRLAFNLV DPILGAEFFY TDGSCNKQSK EGKAGYVTD
q9dsk7 -----

1251 -----
q76053 -----
q90066 -----
q77162 -----
q77249 -----
q73194 GRDKVKLEQ TTNOQAELEA FAMAALDSSG KANIIVDSQY VMGI VAGQPT
q78654 -----
q76630 GRDKVKLEQ TTNOQAELEA FAMAALDSSG KANIIVDSQY VMGI VAGQPT
q9ytuo GRDKVKLEQ TTNOQAELEA FAMAALDSSG KANIIVDSQY VMGI VAGQPT
q78655 -----
q89928 GRDKVKLEQ TTNOQAELEA FAMAALDSSG KANIIVDSQY VMGI VAGQPT
q91b19 GRDKVKLEQ TTNOQAELEA FAMAALDSSG KANIIVDSQY VMGI VAGQPT
q9dsk7 -----

1301 -----
q76053 -----
q90066 -----
q77162 -----
q77249 -----
q73194 ESENRIVNOI IEEMIKKEAI YVAVVPAHKG IGGNQEVNHL VSQGIROQLF
q78654 -----
q76630 ESENRIVNOI IEEMIKKEAI YVAVVPAHKG IGGNQEVNHL VSQGIROQLF
q9ytuo ESENRIVNOI IEEMIKKEAI YVAVVPAHKG IGGNQEVNHL VSQGIROQLF
q78655 -----
q89928 ESENRIVNOI IEEMIKKEAI YVAVVPAHKG IGGNQEVNHL VSQGIROQLF
q91b19 ESENRIVNOI IEEMIKKEAI YVAVVPAHKG IGGNQEVNHL VSQGIROQLF
q9dsk7 -----

1351 -----
q76053 -----
q90066 -----
q77162 -----
q77249 -----
q73194 ESENRIVNOI IEEMIKKEAI YVAVVPAHKG IGGNQEVNHL VSQGIROQLF
q78654 -----
q76630 ESENRIVNOI IEEMIKKEAI YVAVVPAHKG IGGNQEVNHL VSQGIROQLF
q9ytuo ESENRIVNOI IEEMIKKEAI YVAVVPAHKG IGGNQEVNHL VSQGIROQLF
q78655 -----
q89928 ESENRIVNOI IEEMIKKEAI YVAVVPAHKG IGGNQEVNHL VSQGIROQLF
q91b19 ESENRIVNOI IEEMIKKEAI YVAVVPAHKG IGGNQEVNHL VSQGIROQLF
q9dsk7 -----

1401 -----
q76053 -----
q90066 -----
q77162 -----
q77249 -----
q73194 LKIEPAQEE HEKYHNNVKE LVHKGIPOL VAKQIVNSYH SCQKGEALH
q78654 -----
q76630 LKIEPAQEE HEKYHNNVKE LVHKGIPOL VAKQIVNSYH SCQKGEALH
q9ytuo LKIEPAQEE HEKYHNNVKE LVHKGIPOL VAKQIVNSYH SCQKGEALH
q78655 -----
q89928 LKIEPAQEE HEKYHNNVKE LVHKGIPOL VAKQIVNSYH SCQKGEALH
q91b19 LKIEPAQEE HEKYHNNVKE LVHKGIPOL VAKQIVNSYH SCQKGEALH
q9dsk7 -----

1451 -----
q76053 -----
q90066 -----
q77162 -----
q77249 -----
q73194 LKIEPAQEE HEKYHNNVKE LVHKGIPOL VAKQIVNSYH SCQKGEALH
q78654 -----
q76630 LKIEPAQEE HEKYHNNVKE LVHKGIPOL VAKQIVNSYH SCQKGEALH
q9ytuo LKIEPAQEE HEKYHNNVKE LVHKGIPOL VAKQIVNSYH SCQKGEALH
q78655 -----
q89928 LKIEPAQEE HEKYHNNVKE LVHKGIPOL VAKQIVNSYH SCQKGEALH
q91b19 LKIEPAQEE HEKYHNNVKE LVHKGIPOL VAKQIVNSYH SCQKGEALH
q9dsk7 -----

1501 -----
q76053 -----
q90066 -----
q77162 -----
q77249 -----
q73194 LKIEPAQEE HEKYHNNVKE LVHKGIPOL VAKQIVNSYH SCQKGEALH
q78654 -----
q76630 LKIEPAQEE HEKYHNNVKE LVHKGIPOL VAKQIVNSYH SCQKGEALH
q9ytuo LKIEPAQEE HEKYHNNVKE LVHKGIPOL VAKQIVNSYH SCQKGEALH
q78655 -----
q89928 LKIEPAQEE HEKYHNNVKE LVHKGIPOL VAKQIVNSYH SCQKGEALH
q91b19 LKIEPAQEE HEKYHNNVKE LVHKGIPOL VAKQIVNSYH SCQKGEALH
q9dsk7 -----

```

q76630	LKLSRWPIT	HLHTDNGANF	TQOEVKMAVW	WIGIEQSEGV	PYNPQSGVV
q9ylu0	LKLSRWPIT	HLHTDNGVNF	TQOEVKMAVW	WVWIEQSEGV	PYNPQSGVV
q78655	-----	-----	-----	-----	-----
q89928	LKLSRWPIT	HLHTDNGANF	TQODVKMAVW	WIGIEQTEGV	PYNPQSGVV
q91b19	LKLAGRWPIT	HLHTDNGANF	TQOEVKMAVW	WVWIEQAEV	PYNPQSGVV
q9dsk7	-----	-----	-----	-----	-----
1501	-----	-----	-----	-----	1550
q76053	-----	-----	-----	-----	-----
q90066	-----	-----	-----	-----	-----
q77162	-----	-----	-----	-----	-----
q77249	-----	-----	-----	-----	-----
o73194	EAMNHHLKQ	ISRIREQANT	VETIVLMAVH	CMNFKRRGGI	GMTTPSERLI
q78654	-----	-----	-----	-----	-----
q76630	EAMNHHLKQ	ISRIREQANT	VETIVLMAVH	CMNFKRRGGI	GMTTPAERLI
q9ylu0	EAMNHHLKQ	ISRIREQANT	VETIVLMAVH	CMNFKRRGGI	GMTTPAERLI
q78655	-----	-----	-----	-----	-----
q89928	EAMNHHLKQ	IDRIREQAVS	IETIVLMATH	CMNFKRRGGI	GMTTPAERIV
q91b19	ESMNHHLKQ	IDKIRDOANS	IETIVLMAVH	CMNFKRRGGI	GMTTPAERII
q9dsk7	-----	-----	-----	-----	-----
1551	-----	-----	-----	-----	1600
q76053	-----	-----	-----	-----	-----
q90066	-----	-----	-----	-----	-----
q77162	-----	-----	-----	-----	-----
q77249	-----	-----	-----	-----	-----
o73194	NMITTEQEIQ	FLQAKNSKLK	NFRVYPREGR	DQMWKPGGEL	LWKGDGAVIV
q78654	-----	-----	-----	-----	-----
q76630	NMVTAEQEIQ	FLQAKNSKLQ	NFRVYPREGR	DQMWKPGGEL	LWKGDGAVIV
q9ylu0	NMISTEQEIQ	FLQTKNLKFK	NFPVYPREGR	DQMWKPGGEL	LWKGDGAVIV
q78655	-----	-----	-----	-----	-----
q89928	NMITTEQEIQ	FLQTKNLKFK	NFRVYPREGR	DQMWKPGGDL	LWKGEQAVII
q91b19	NMITTEQEIQ	FOQTKNSKFK	NFRVYPREGR	DQMWKPGGEL	LWKGEQAVII
q9dsk7	-----	-----	-----	-----	-----
1601	-----	-----	-----	-----	1645
q76053	-----	-----	-----	-----	-----
q90066	-----	-----	-----	-----	-----
q77162	-----	-----	-----	-----	-----
q77249	-----	-----	-----	-----	-----
o73194	KVGTDIKIIP	RRKAKIIRDY	GGQELDSSS	HLEGARENGE	VA---
q78654	-----	-----	-----	-----	-----
q76630	KVGADIKIIP	RRKAKIKDY	GGQEMDSCS	NLEGAREDEGE	VA---
q9ylu0	KVGTDIKVVP	RRKAKIIRDY	GGQELDSCP	HLEGAREDEGE	VA---
q78655	-----	-----	-----	-----	-----
q89928	KVGTEIKVIP	RRKAKIIRNY	GGGKELDCSA	DVEDTMOARE	VAQSN
q91b19	KVGTEIKVVP	RRKAKIKDY	GGGKEVDSST	NMEDTROTGE	VA---
q9dsk7	-----	-----	-----	-----	-----